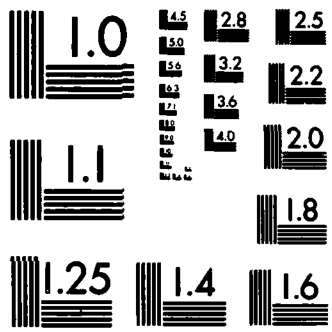
AD-A122 285 PHASE I EVALUATION OF TRASANA TECHNICAL STAFF PRODUCTIVITY MEASUREMENT SY..(U) ARMY TRADOC SYSTEMS ANALYSIS ACTIVITY WHITE SANDS MISSILE RAN.. JAN 82 1/2

UNCLASSIFIED TRASANA-TR-5-82 SBI-AD-F050 072 F/G 5/9 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADF050072

AD A122285

TRASANA

TECHNICAL REPORT 5-82

PHASE I EVALUATION OF TRASANA TECHNICAL STAFF PRODUCTIVITY MEASUREMENT SYSTEM TEST (TPMS-TE)

JANUARY 1982

Approved for public release; distribution unlimited.

DEPARTMENT OF THE ARMY
US ARMY TRADOC SYSTEMS ANALYSIS ACTIVITY
WHITE SANDS MISSILE RANGE
NEW MEXICO 88002

DISCLAIMER

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other official documentation.

WARNING

Information and data contained herein are based on the input available at the time of preparation. The results are subject to change.

EXECUTIVE SUMMARY

This report presents the results of the Phase I (the first scoring year) evaluation of the three-year test implementation of the US Army TRADOC Systems Analysis Activity (TRASANA) Technical Staff Productivity Measurement System (TPMS).

TPMS is a methodology designed to measure the productivity of approximately 250 technical staff members who are mostly scientists and engineers engaged in various facets of operations research and applied problem solving. Most of the work is performed by teams organized to meet the requirements of each project. The organization completes 60 or more projects a year which are quite diverse in nature. Each project is required by TPMS to have a project plan, which is a detailed planning document including estimates of manpower and computer resources needed for the project, as well as a time schedule for the project.

The project plan is submitted for approval to the Productivity Measurement Board (PMB) which is responsible for monitoring resource estimates and schedules. Once the plan is approved, the team executes it. When the project is completed, the project is assigned scores based on how well the actual resource utilization and time schedule agreed with the plan, as well as a product quality score. The product quality score is assigned by a quality control group called the Product Review Board (PRB). The scores for all of the factors are multiplied together to arrive at the Project Score (PS). Two or more PS's can be combined and weighted by expended manpower to form the Productivity Index (PI). PI's are calculated for each individual and for organizational units of TRASANA.

TPMS was developed at TRASANA and adopted for a three year trial implementation starting 1 Oct 80. The first scoring period was a truncated year (nine months) ending 30 Jun 81 to put the system on the 1 July to 30 June year.

The objectives of this study were (1) to determine how the TRASANA technical staff (TTS) perceives TPMS, (2) to determine the degree to which TPMS is a meaningful measure of productivity, and (3) to develop a TPMS data base consisting of Phase I TPMS-T results.

The first study objective was answered by surveying the TTS at the end of the first year. The second objective was answered by an in-depth analysis of the first year TPMS scoring results, a mathematical analysis of the scoring algorithm, a functional analysis of the PMB and PRB, and by a PI validation analysis. The PI validation consisted of developing a measure of productivity which was independent of TPMS, and which could be used for comparison with the PI. This independent measure, or criterion measure (CM), consisted of confidential supervisory ratings obtained by personal interview. The third objective was achieved by forming a computer data file consisting of both the survey and first year scoring results.

The major findings were:

- The staff reaction to TPMS was mixed. They see TPMS as a potentially useful managerial tool, but there is a high level of concern about PI inaccuracies, system manipulation, and the effect on the creative and cooperative atmosphere at TRASANA.

- The system was not seen as fair by the staff.

- TPMS reflects organizational productivity in the classical sense (value of output over value of input, and captures both efficiency and effectiveness) but its accuracy needs to be improved.

- The resource estimates, especially manpower, were not adequately controlled by the PMB, which decreased the accuracy of the system.

- The individual PI did not accurately reflect individual productivity.

- There is considerable potential for input effort/output score mismatch with the individual PI.

- The scoring algorithm, especially the effect of judgmental and discretionary awards, was not well understood.

- Organizational units with an overhead mission are disadvantaged by including overhead (activities that are not project specific) in the PI computation.

- The organizational PI was more accurate than the individual PI, but it will be much more accurate if the resource estimates are controlled and overhead is more consistently reported across organizational units.

- The PRB serves a valuable quality control function, but the quality scoring was much less successful.

The major conclusions were:

- TRASANA is a healthy organization with high morale and a motivated work force.

- TPMS is a conceptually sound productivity measurement system, but many of its functional elements need revision.

- The individual PI should be eliminated.

- The organizational PI would be strengthened by eliminating overhead from the PI calculation.

- A new method of controlling resource estimates is required.

- The scoring algorithm is unnecessarily complicated.
- Firm guidelines are needed for overhead accounting.

Major recommendations included:

- Use only three scoring factors - manpower, timeliness and quality - and score all three by categories, e.g., high, medium, and low. The three category scores for manpower and for timeliness should reward accuracy in planning.
- Change the 125% rule (percentage of available manpower committed to projects) to the 110% rule.
- Institute an after-the-fact division level check with reward for accuracy in estimating manpower and timeliness.
- Eliminate the individual PI.
- Establish overhead charging guidelines with a specified percentage of manpower allowed to be charged to overhead without penalty.
- Eliminate overhead from organizational PI computation.
- Establish a scoring checklist (with scoring criteria) to be used by the PRB for scoring each product.
- Assign a productivity rating to the PRB.
- Maintain management intervention to avoid penalties in PI from undertaking risky projects.

An alternative TPMS scoring procedure incorporating the above recommendations is included in the report.

TABLE OF CONTENTS

PHASE I EVALUATION OF TRASANA TECHNICAL STAFF
PRODUCTIVITY MEASUREMENT SYSTEM-TEST (TPMS-TE)

CHAPTER 1

INTRODUCTION

1.1 PURPOSE

This report presents the results of the Phase I (the first scoring year) evaluation of the three-year test implementation of the US Army TRADOC Systems Analysis Activity (TRASANA) Technical Staff Productivity Measurement System (TPMS).

1.2 BACKGROUND

a. TPMS is a methodology designed to measure the productivity of the TRASANA technical staff (TTS), its organizational elements, and its individual employees. It yields a single dimensionless index as the measure of productivity. The formulation of this index is intended to be consistent with the traditional definition of productivity as value of output divided by value of input. In principle, TPMS places TRASANA on a "fixed price plus incentive" basis for producing its various products. TPMS was designed with the following attributes in mind:

(1) Relevant - to the organization, its products, and to the jobs that produce the products.

(2) Rational and fair - so that it can be accepted and used by a sophisticated workforce.

(3) Affordable - in terms of direct and indirect/overhead costs for operation and maintenance.

(4) Flexible - to couple with a dynamic workload, organization, and operations.

(5) Simple - both in principle and practice.

(6) Transparent - so as to facilitate understanding of cause and effect relationships.

(7) Durable - to support multi-year use.

b. TPMS was adopted at TRASANA for a three year trial beginning 1 Oct 80. The first scoring period ended 30 Jun 81, and was nine months in length. Subsequent years will run from 1 July through 30 June. Management formed a small internal evaluation team in Mar 81, and charged the team with developing and executing an objective evaluation of the TPMS test implementation. The study plan developed for evaluating the initial implementation period is in appendix A. In keeping with management's desire to have as independent and

objective evaluation as possible, the study plan was developed independently by the study team, and reviewed, but not changed, by management. The evaluation effort was directed toward determining the extent to which TPMS met its design attributes.

1.3 THE SYSTEM

A complete description of TPMS, its development and the three year trial implementation period can be found in other documents.[1] This brief overview is provided for the reader who is not familiar with the system.

a. TRASANA has nearly 250 personnel who are referred to as analysts or technical staff. They are mostly scientists and engineers who engage in various facets of operations research and applied problem solving. The *modus operandi* is for a team to be assigned to a given project. The team is responsible for clearly defining the problem, developing and executing an appropriate methodology, and publishing a report documenting the results of the effort.

b. TPMS basically provides a method for assigning a numerical score to each project after it is completed. This score is called the project score (PS). The PS is converted to a weighted index (weighted by invested resources) which can in turn be converted to a weighted productivity index (PI) for individuals as well as organizational units.

c. The mechanics of the system are rather straight forward. When a project is identified, a senior analyst is assigned as the study manager (SM). The SM is responsible for negotiating details of the project with any outside agencies and writing a project plan. The project plan is a detailed planning document which breaks down the project into tasks and subtasks, specifies the manpower and computer resources needed for each subtask, and also specifies what the product of each subtask will be and when it will be completed. Once this document is completed, it is submitted to the Productivity Measurement Board (PMB), which consists of the Assistant Deputy for Technical Operations (chairperson) and the eight functional division chiefs. The PMB reviews each project plan for completeness and to insure the resource requirements and completion dates are realistic.

d. Once a project plan is approved by the PMB, the basis for assigning a project score has been established. There are five factors contributing to the PS: manpower utilization efficiency (R), computer utilization efficiency (C), timeliness (T), product quality (Q), and priority (P). These factors each take on a limited value as shown in table 1-1. The R and C factor scores are essentially a ratio of planned to expended resources. T is based on how

---

[1] *TRASANA Memorandum 5-4, TRASANA Technical Staff Productivity Measurement System (TPMS)-Test, August 1980.*
TRASANA Technical Documentation 37-80, Productivity Measurement System Design and Development, December 1980.

TABLE 1-1. TPMS SCORING FACTOR VALUES

FACTOR	MINIMUM VALUE	MAXIMUM VALUE
R	0.7	1.2
C	0.9	1.1
T	0.8	1.1
Q	1.0	1.5
P	1.0	1.2

close the product is completed to the projected completion date. Q is assigned by the Product Review Board (PRB), which is an internal quality control mechanism consisting of three permanent (voting) members augmented by additional analysts selected on an *ad hoc* (non-voting) basis. P is 1.2 if the study is conducted for an external customer, and is negotiated with the PMB in other cases. The PS is computed by multiplying all of the factor scores together:

$$PS = R \times C \times T \times Q \times P$$

1.4 PROBLEM

The problem addressed in the Phase I evaluation was to determine the degree to which TPMS provides a meaningful measure of the productivity of the TRASANA technical staff and its organizational elements.

1.5 IMPACT OF PROBLEM

TPMS is a new system. Numerous decisions were made in designing the system which involved trade-offs. The only way to know how well the final result achieves its goals is to test it. Furthermore, a comprehensive evaluation of the test is required to fully capture the effects of the system, and to document its strengths and weaknesses. Failure to comprehensively evaluate TPMS performance during the test period could result in implementing a system which does not adequately accomplish its objective, or which even may be counterproductive.

1.6 OBJECTIVES AND STUDY QUESTIONS

a. To determine how the TRASANA technical staff perceives TPMS. The study questions for this objective are:

(1) What is the TTS general attitude toward TPMS?

(2) What is the perceived motivational impact of TPMS?

(3) Is TPMS seen as a viable managerial tool?

b. To determine the degree to which TPMS is a meaningful measure of productivity. The study questions for this objective are:

(1) To what extent do TPMS ratings agree with confidential supervisory productivity assessments for individuals and branches?

(2) Does TPMS have generally the same impact on the different TRASANA organizational divisions?

(3) What are the functional relationships among the scoring algorithm variables?

(4) What is the operational impact of the Productivity Measurement Board (PMB) and Product Review Board (PRB)?

c. To develop a TPMS data base consisting of Phase I TPMS-T results. Available data on resource implications will be included. This data base will form the benchmark for comparison of future results.

1.7 SCOPE

a. The Phase I evaluation was limited to the results of the first trial implementation year (nine months) of TPMS.

b. This evaluation was limited exclusively to TPMS. While TPMS was developed in part as a local response to requirements of the new Performance Appraisal (PA) and Merit Pay (MP) systems resulting from the 1978 Civil Service Reform Act, no assessment of the PA or MP systems was intended nor attempted.

c. TPMS-TE is an assessment of a specific productivity measurement system. It is not intended to be a comprehensive study of methods to measure TRASANA productivity or research and development type productivity in general.

d. One of the desired characteristics of TPMS is that it be affordable. Available data will be collected on resources expended on the first year test of TPMS; however, judgments on whether the costs are a worthwhile investment are beyond the scope of this study.

1.8 ASSUMPTIONS, CONSTRAINTS, AND LIMITATIONS

a. TPMS will be applied equally to all TRASANA Technical Staff (military and civilian).

b. Key organizational supervisors know the relative long term productivity of employees they supervise.

c. Division chiefs, branch chiefs, and the TTS provided complete and accurate information for the study.

d. Not all of the TTS worked on TPMS scored projects during the first scoring period.

e. Since there was no uniform, quantitative productivity measure prior to TPMS, there is no basis for comparing test period productivity with past productivity.

f. There is no independent measure of productivity with which to directly compare the TPMS results.

CHAPTER 2

METHODOLOGY

2.1 APPROACH

The evaluation of the functioning of TPMS during the first trial scoring period was approached in three ways as depicted in figure 2-1. First, the reaction of the TRASANA technical staff was assessed. Data for this assessment were collected by survey at the end of the first scoring period, but before scoring results were generally available. Second, an independent measure (criterion) was developed to determine the degree to which the productivity index (PI) reflects productivity. Third, the functioning of selected TPMS components were evaluated using data collected from several sources. The procedures utilized for each of these analyses are presented in the following paragraphs.

2.2 TRASANA TECHNICAL STAFF PERCEPTIONS

2.2.1 TRASANA Technical Staff (TTS) Survey

a. The survey was developed to answer specific management questions, as well as to provide information required by the study team for the evaluation. The results of an invited Organizational Effectiveness Visit, 9-12 Feb 81 indicated considerable uneasiness about TPMS.[1] Management desired more precise information about the level of concern among the staff, and wanted to know what particular aspects of TPMS were precipitating the anxiety. In providing this information for management, a baseline would also be established which could be used to document any shifts that occurred over the remaining two years of the trial implementation.

b. In addition to staff reaction to TPMS, the survey provided the following types of information:

- Respondent background (survey was anonymous)
- Feelings toward TRASANA
- Job satisfaction
- Understanding of TPMS

c. After specific objectives for the survey were defined, questions were constructed, and a draft questionnaire was printed and pilot-tested. The pilot test consisted of administering the survey to two groups (n = 5, 6) of TRASANA analysts (cross-validation design). The survey was given to each

---

[1] *TRADOC OE Team (ODCS PAL) Memorandum for Director, TRASANA, 24 Feb 81, subject: Organizational Effectiveness Visit, 9-12 Feb 81.*

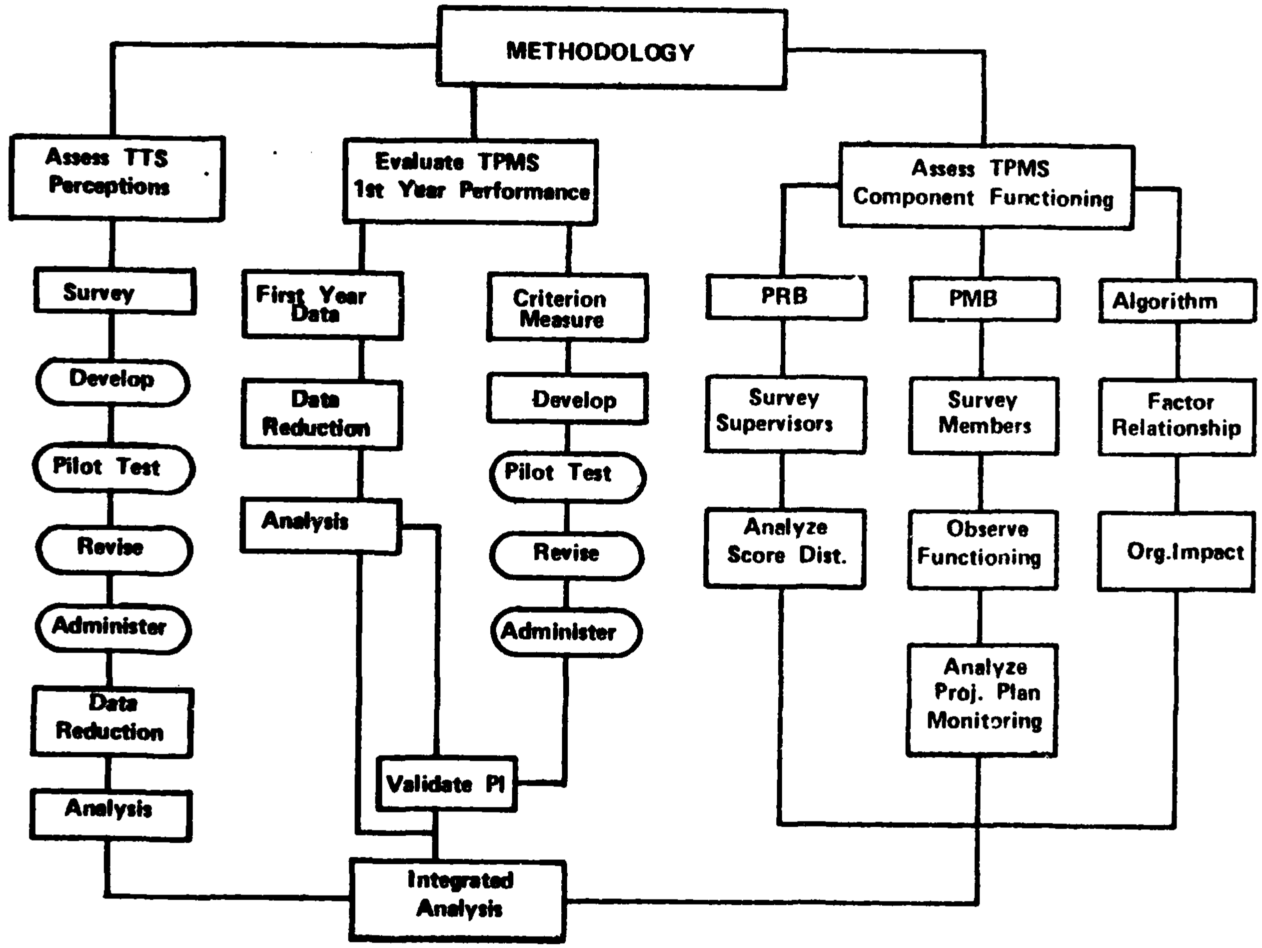


Figure 2-1. TPMS-TE Methodolgy

analyst individually, and comments were solicited on clarity, difficulty in answering any questions, and whether the questions provided ample opportunity for expressing all feelings and concerns about TPMS. Based on these comments and a statistical item-analysis, revisions to the survey were made. The revised survey contained 91 items organized into three parts: demographics, attitude toward TRASANA and TPMS, and TPMS familiarity. There was also a comment page attached to the survey. A copy of the survey is in appendix B.

2.2.2 Sample

a. The survey was administered to 190 analysts, or approximately 80 percent of the TTS with experience under TPMS. Those who did not take the survey fell into one or more of several categories. The survey authors and those who took the pilot survey were not considered part of the available population. Furthermore, persons who were identified as new to the organization, and therefore not familiar with either the organization or TPMS, were also excluded. The majority of the remaining persons who did not take the survey were unavailable due to a combination of leave and work related travel requirements.

b. Table 2-1 provides a comparison of the survey sample and the TTS by grade. Examination of the data reveals the relative homogeneity of the two groups. The two distributions were compared statistically and found not to be different.[2] The sample is therefore considered statistically representative of the TTS by grade.

c. Table 2-2 shows the participation rate in the survey by division. Percent participation in the table is not necessarily percent of total assigned who took the survey, because pilot test participants, evaluation team members, and identified new personnel were excluded. All division chiefs and 20 branch chiefs took the survey. A statistical comparison was made between the distribution of survey respondents by division, and the actual assigned personnel by division to determine if the division representation in the survey sample was biased.[3] There was no statistically significant difference in the distributions; therefore, no division was over or under represented to the point of biasing the results.

---

[2] *The test for independence was with the Chi-square technique.* $X^2 = 11.55$, $df = 8$, *which is not statistically significant at the* $p < .05$ *level* $(X^2_{8,.05} = 15.51)$.

[3] *This comparison was by Chi-Square using actual staffing figures from the Aug 81 TRASANA Manpower Account Roster.* $X^2 = 6.19$, $df = 7$, *which is not statistically significant at the* $p < .05$ *level* $(X^2_{7,.05} = 14.07)$. *Table C-1 in appendix C shows the composition of the TTS and the survey sample by division.*

TABLE 2-1

COMPARISON OF COMPOSITION OF SURVEY

SAMPLE AND TECHNICAL STAFF POPULATION BY GRADE

GRADE	NUMBER IN SAMPLE	PERCENT OF SAMPLE	PERCENT OF POPULATION[a]
CIVILIAN:			
GS-7	3	1.6	4.2
9	9	4.7	3.5
11	14	7.4	6.5
12	51	26.8	25.4
13	53	27.9	26.2
14	27	14.2	12.3
15	8	4.2	3.5
Other	3	1.6	--
MILITARY:			
Officer	16	8.4	12.7
Enlisted	6	3.2	5.8
TOTAL	190	100	100.1[b]

[a]*Source: TRASANA Manpower Management Branch, August 1981 Manpower Account Roster.*

[b]*Does not total 100% due to rounding.*

TABLE 2-2

PARTICIPATION IN SURVEY BY ACTIVITY AND DIVISION

UNIT	PARTICIPTION (PERCENT)
DIVISION:	
A	64.5
B	83.3
C	72.7
D	86.5
E	88.9
F	62.5
G	75.0
H	97.3
ACTIVITY:	80.3

2.2.3 Data Collection Procedures

a. The survey was administered to the TTS on 29 and 30 June 1981. The timing was selected to allow maximum exposure to the system for familiarity, but yet to gather perceptions before the first scoring period results were available. Since the first scoring period ended 30 June 1981, the survey captured perceptions at the end of the period, during which 61 projects were completed and scored. Makeup sessions were held on 20 and 24 July.

b. A letter signed by the Assistant Deputy for Technical Operations was sent to each of the eight division chiefs approximately one week before the scheduled administration date. The division chiefs were all aware that the survey was to be administered at the end of the first scoring year because they had reviewed and approved the project plan in the Productivity Measurement Board (PMB) meeting (12 June 1981). At the TRASANA staff meeting of division chiefs on 25 June, maximum participation in the survey was encouraged. Division secretaries were also phoned on administration dates as a reminder.

c. The survey was scheduled to provide as little work interruption as possible, and to be as convenient as possible while still under controlled conditions. Four three-hour time blocks were established over the two days with two divisions being scheduled for each time slot. Analysts could come to the scheduled conference room at their convenience, complete the survey, and leave. The time schedule was only to avoid congestion. No attempt was made at schedule enforcement, and all requests to take the survey in a different time slot were honored. The survey took 45 minutes to an hour, and was taken anonymously. Group administration was employed to control for discussion of questions and answers while responding and to achieve a more representative sampling than could be achieved by a mailing procedure.

2.3 EVALUATION OF TPMS FIRST YEAR PERFORMANCE

The evaluation of the TPMS first year performance involved the following steps:

- gathering and analyzing first year scoring results
- conducting a literature survey on productivity evaluation
- developing a productivity criterion measure (CM)
- collecting CM data
- PI evaluation analysis

The methodology employed for each of these steps is discussed separately.

2.3.1 First Year Scoring Results

a. TRASANA Memo 5-4, describes the TPMS three year implementation, scoring procedures, and record keeping requirements. Essentially, each of the eight TRASANA divisions are responsible for keeping records of their projects, and for reporting the results at the end of each scoring year.

b. A set of three forms was developed which would provide the necessary information from each division for this evaluation, as well as provide a convenient summary record which would meet the needs of the divisions.[4] A copy of this set of forms, TPMS Division Scoring Sheet, TPMS Branch Scoring Sheet, and TPMS Individual Scoring Sheet is in appendix D.

c. The forms were distributed to the division chiefs at a PMB meeting with instructions to furnish a copy to the TPMS-TE evaluation team not later than 15 August 1981, which was enough time for all scoring records to be available.

d. Project scoring data were obtained from the TPMS scored project log, which is maintained by the PMB chairman.

e. After all scoring data were received, a descriptive summary was prepared and analyzed for each of the scoring variables, for each project, and for each scoring unit. These data formed the Phase I TPMS-T data base.

---

[4] *The final version forms were a slighty modified version of those developed by ATAA-TE.*

2.3.2 Literature Survey Summary

A comprehensive literature review was beyond the scope of this study. Instead, an attempt was made to select the most relevant works as a framework for this evaluation. The works selected for review fell into four general categories:

- documents pertaining specifically to TPMS
- Research and Development (R&D) productivity measurement
- performance appraisal (PA)
- peer ratings

All available documentation on TPMS were reviewed for background, and the best available references on measuring productivity of R&D or scientist and engineer organizations were consulted for background in developing the CM. References on PA and peer ratings were consulted as these techniques became important in developing the CM. A detailed summary of the works consulted is not included, because excellent literature summaries appear elsewhere.[5] Conclusions from the literature survey are outlined to show the reasoning behind the specific study procedures employed.

a. Measuring R&D Productivity. The predominant approach to measuring productivity among scientists and engineers is to opt for some easily countable measure such as number of publications (Blose & Stansel, 1975; Katzen, 1975; Koser, 1976; Newburn, 1972; Technical Report 74-3, USAF AFIT, 1974; and Technical Report 78-1, USAF AFIT, 1978). One ambitious attempt to compare the knowledge production of major universities failed to find a more acceptable measure than publications (Clark & Guba, 1977). The obvious problem with counting products is that the quality of the products is not considered. One high quality product may be equal to several lesser quality works. Edwards and McCarrey (1973), after an extensive literature review concluded that more study is needed on how to measure the performance of research or knowledge type workers. The difficulty is measuring quality or relevant worth. The most credible method of doing this is by blind peer review. This method involves presenting a small group of widely recognized professionals in a specific field with works not identified by author to be rated according to specified rating criteria. This "blind review" is the method used by many professional journals to select articles for publication (commonly referred to as "refereed" journals).

---

[5] *The works consulted are arranged by subject area in the bibliography. The most complete general reference is R&D Productivity published by the Hughes Aircraft Company (1978). Other references listed in the bibliography containing relevant literature summaries include: Christ (1981), Edwards and McCarrey (1973), Schainblatt (1981), and Technical Report 79-9 (USAF HRL, 1979).*

b. Peer Ratings. The most frequent use of peer ratings reported in the literature is in training courses. Very little was found that related to scientists and engineers. No evidence was found to support the use of peer ratings for determining product quality in an organization like TRASANA. Indeed, the problems that are reported in attempts to use peer ratings suggest questionable validity. It was anticipated that with the diverse backgrounds of the TTS, documented problems with peer ratings would be magnified.

c. Performance Appraisals. Once it became evident that peer ratings were not a viable alternative for use as a CM, it was apparent that some form of supervisor ratings or evaluations would be necessary. The PA literature was searched to find the most appropriate factors to include in the appraisals as well as the best structure or format. From the literature it was found that the most reliable and valid appraisal is a global or overall rating. This method was adopted for the CM.

2.3.3 Criterion Measure

a. Official employee performance appraisals were not adequate for the CM because of the well known problems such as halo effect, rater bias, and variation in rater standards. The most accurate supervisor judgments were required, and these judgments had to be quantitative to allow comparison with employee PI rankings.

b. The selected method involved asking first line supervisors (branch chiefs) to provide a confidential rank order of their employees. This ranking was done in a private interview. Management provided the guarantee that this information would remain confidential. Neither superiors nor subordinates had access to this information. To preclude any chance of compromise, the confidential information from branch chiefs was destroyed as soon as the necessary data pairings for the analysis were completed. In addition to the ranking of employees, a rating was also obtained. The rating was an overall performance evaluation on a scale of 1 to 10, with 5 being average for TRASANA, and 10 being the most outstanding. These two confidential data elements for each individual, rank order in branch and overall job performance rating, became the productivity CM for the PI evaluation.

2.4 TPMS COMPONENT FUNCTIONING

This portion of the evaluation effort consisted of determining how well the Productivity Measurement Board (PMB), the Product Review Board (PRB), and the TPMS scoring algorithm functioned when considered all together.

2.4.1 PMB

Data were collected from three sources to determine how well the PMB worked the first year. First, a meeting was observed. Second, the PMB members were interviewed. Third, the first year scoring results were analyzed to determine if the PMB monitoring of project plans was achieving its objective of keeping manpower, computer, and timeliness estimates from being inflated.

2.4.2 PRB

The PRB functioning was determined by interviewing selected PRB members, and division chiefs, branch chiefs, and study managers who experienced PRB scoring during the first scoring year. The scores given by the PRB during the first scoring period were also analyzed.

2.4.3 Algorithm

The TPMS scoring algorithm was analyzed mathematically to clearly identify factor relationships.

CHAPTER 3

FINDINGS: TECHNICAL STAFF SURVEY RESULTS

This chapter presents the results of the survey administered to the TRASANA Technical Staff (TTS) at the end of the first scoring year to achieve the first study objective: to determine how the TTS perceives TPMS. The following study questions for this objective are answered:

- What is the TTS general attitude toward TPMS?

- What is the perceived motivational impact of TPMS?

- Is TPMS seen as a viable managerial tool?

Emphasis was on quantifying the general attitude toward TPMS, determining its perceived motivational impact, and assessing the perceptions of TPMS as a viable management tool. Since the general predisposition of the TTS would be an important factor in how a new productivity measurement system would be accepted, organizational climate data were also collected in the survey, and are reported in this chapter.

3.1 THE SURVEY

a. A complete copy of the survey is in appendix B. It consists of 91 items plus a comment page. The first 14 items pertain to background information. The data obtained from these items are summarized in tables C-2 through C-7 of appendix C. These data document the fact that the TTS is a highly educated, experienced, stable work force consisting primarily of scientists and engineers.

b. Seventy-five items were constructed on a six-point forced-choice Likert scale with no neutral response category. The Likert scale was selected because it gives not only an agree or disagree response to each item, but also yields an intensity, or degree of agreement or disagreement, which is much more revealing than a two category (yes or no) response. Sixty-two items were grouped into the following clusters:

- Attitude toward TRASANA

- Job Satisfaction

- Control of work

- General reaction to TPMS

- Fairness of TPMS

- Understanding of TPMS

- Usefulness of TPMS as a management tool

The additional 13 items did not fall into any cluster, but addressed specific points of interest.

3.2 SURVEY RESULTS

The results of the survey clusters are discussed in this section. Even though cluster scores are more stable and contain less error variance than individual item responses, it is recognized that review of the distribution of responses to individual items may be of interest to some readers. Tables C-8 and C-9 in appendix C give the response distributions for all of the questions in the survey except the demographic items, which are in tables C-2 through C-7.

3.2.1 Cluster Means and Reliabilities

a. The results of the survey clusters are in table 3-1. The reliability coefficients (coefficient alpha) range from 0.70 to 0.86 for individual clusters, indicating that the items in the clusters were generally a homogeneous set of items (all measuring the same thing). The reliability coefficient of 0.89 for the entire 75 items is high for a survey of this type.

b. The results were scored so that the higher the score, the more positive the attitude. The mid-point of the six-point scale is 3.5. Scores falling in the 3.0 to 4.0 range are neither clearly positive nor negative due

TABLE 3-1
TTS SURVEY CLUSTER (SUBSCALE) RESULTS

CLUSTER	NUMBER OF QUESTIONS	RELIABILITY COEFFICIENT	MEAN	STANDARD DEVIATION	SE_M
ORGANIZATION:					
Attitude toward TRASANA	12	.86	4.32	.84	.31
Job Satisfaction	9	.81	4.51	.89	.39
Control of work	3	.70	4.63	1.08	.59
TPMS:					
General Reaction	11	.73	3.45	.76	.39
Fairness	4	.82	2.96	1.21	.51
Understanding	11	.76	4.00	.87	.21
Management Tool	12	.70	3.46	.72	.39
CLUSTER TOTAL	62	.88	3.82	.72	.25
SURVEY TOTAL	75	.89	3.80	.45	.15

to possible measurement error.[1] The three clusters measuring organizational climate are clearly in the positive range, showing a very favorable regard for the organization, job satisfaction, and control of work (important for R&D organizations). The TPMS clusters are less positive. The general reaction and usefulness as a management tool are clearly in the neutral range. The staff generally considers itself to understand TPMS, but does not think TPMS is fair.

3.2.2 Attitude Toward TRASANA

Overall, the staff feels that TRASANA is a good place to work. Ninety-four percent of the respondents like working at TRASANA, 90 percent are proud of TRASANA's products, 91 percent think that TRASANA has high standards, and 87 percent feel that people are treated fairly.

3.2.3 Job Satisfaction

a. In general the TTS expressed a high degree of job satisfaction. Sixty-eight percent indicated they get a lot of satisfaction from their job, and 75 percent report their job is challenging. Of those not experiencing much personal satisfaction from their jobs, 43 percent indicated that their skills were not well matched to the requirements of their jobs.

b. Fifty-four percent of the respondents do not feel they get useful feedback on job performance. This is indicative of a problem area. Analysts generally have a lot of freedom in how they do their job (92 percent say they do), and supervisors generally meet with approval, yet over half do not feel their supervisors are doing an effective job of performance appraisal. Meaningful performance evaluation and feedback is important for many reasons, including helping persons to improve themselves through constructive criticism, and maintaining a realistic perspective on future expectations, such as promotion potential. A majority (52 percent) do not see TPMS as being useful for individual appraisals, which means it will not likely fill this void.

3.2.4 General Reaction to TPMS

The aggregate response to this cluster lies near the scale midpoint--neither positive nor negative. This reflects mixed feelings as well as disagreement among the staff. There is considerable concern that too much emphasis will be put on TPMS results to the exclusion of important job performance factors not captured by TPMS scores (60 percent do not believe TPMS emphasizes the most important factors in productivity). On the positive side, 81 percent

---

[1] *Any measurement, physical or psychological, is only accurate to some specified degree, and contains some tolerance or error. A band of confidence can be built around a test score to show the upper and lower bounds of a given score for a specified reliability and distribution. This band of confidence can be calculated by taking a score plus and minus the Standard Error of Measurement (SE_M). The SE_M for this survey, shown in the last column of table 3-1, reflect that the scores are rather stable.*

think TPMS is being given a fair chance, 69 percent believe that manpower resources eventually can be realistically estimated, 61 percent believe computer resources can be realistically estimated with more experience, and 34 percent have seen some positive effects of TPMS (basically in project planning and resource allocation). On the other hand, 67 percent do not think it is easy to understand, 58 percent believe that it is not flexible enough for TRASANA's variety of projects, and 83 percent believe it can be easily manipulated.

3.2.5 Fairness

The response on this cluster was the least positive. Fifty percent said TPMS was not fair (19 percent strongly disagreed with the statement that TPMS is basically a fair system). Sixty-three percent do not believe divisions have an equal chance, 59 percent do not believe it is objective, and 61 percent do not belive that the most productive individuals will get the best productivity scores under TPMS.

3.2.6 TPMS Understanding

Of the clusters pertaining to TPMS, the response to this one was the most positive. It should be noted that this cluster contained items designed to reveal system understanding as well as items asking for staff perceptions of their own and their peer's understanding. The items designed to test system understanding showed the staff could be divided into three approximately equal groups: high, medium, and low understanding. The supervisors had the highest level of understanding, and the newest employees the lowest. Now that the activity has completed a scoring cycle, it is anticipated that understanding of the system has improved considerably since the survey was administered.

3.2.7 Usefulness as a Management Tool

Sixty-four percent indicated they thought TPMS would be a useful management tool. There is general agreement that management practices have changed as a result of TPMS (64 percent) and that more planning is now being done (72 percent), but only 33 percent believe management practices have improved. Fifty-eight percent believe TPMS is a clear change in emphasis by management, even though 62 percent do not think it will motivate people to be more productive.

3.3 GROUP COMPARISONS

3.3.1 Supervisors Versus Nonsupervisors

Mean cluster scores for supervisors and nonsupervisors were compared statistically. In every case supervisors had significantly higher (more positive) scores than the nonsupervisors.[2] Figure 3-1 shows the scoring

---

[2] *These comparisons were made with t-tests. Table C-10 in appendix C gives the group means and standard deviation for each cluster, as well as the t values, which are are all significant beyond the .05 level.*

profile of the two groups. The clusters are shown along the bottom or horizontal axis, and the cluster mean scores are scaled along the vertical axis. This profile analysis shows that supervisors have a consistently more positive attitude than nonsupervisors on both the organizational and TPMS clusters. The fact that the two profiles are almost parallel indicates that the supervisors are not only consistently more positive, but uniformly more positive as well (there are no group-cluster interaction effects). While it is not surprising that supervisors are more positive about the organization, the near perfectly parallel profiles indicates that supervisors and nonsupervisors have the same relative feelings about TPMS.

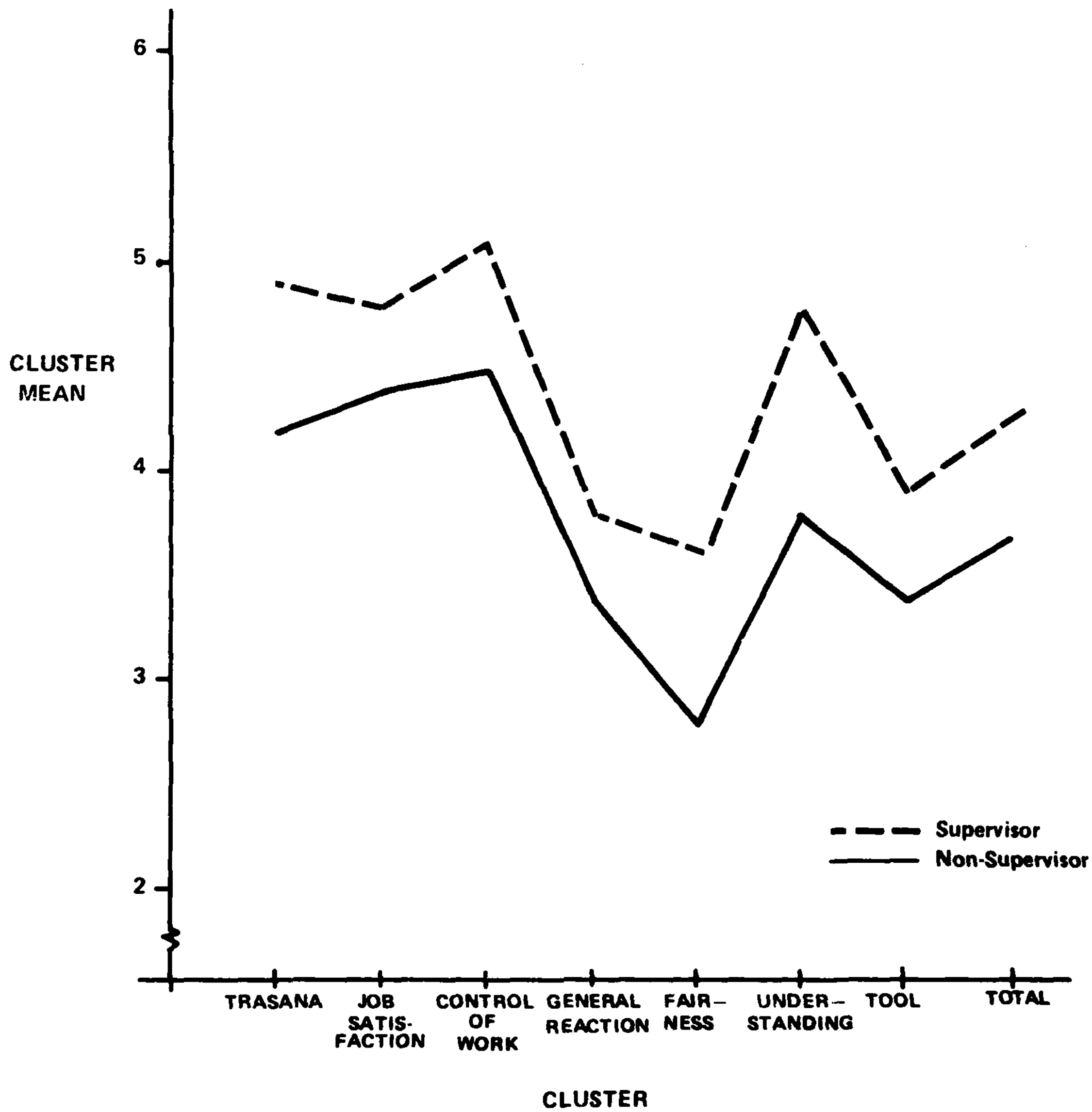


Figure 3-1. Supervisor/Non-Supervisor Profile Analysis

3.3.2 Other Group Comparisons

a. Mean cluster scores were also compared for persons grouped on the following variables:

- Age
- Time in Federal Service
- Time at TRASANA
- Grade
- Military/Civilian
- Division

Most of these comparisons revealed no significant differences between groups, but a few interesting trends were discovered.

b. There was a tendency for those who had been in the organization longer to have a better understanding of the system, to see the system as less unfair, and to have a more positive general reaction to TPMS. It should be recognized that nearly all of the supervisors fall in the grouping or category that has been in the organization over nine years. The same pattern emerged for the grade and time in federal service categories, because these groups are essentially the same people.

c. The military staff members tended to be more positive toward TPMS in general than the civilian staff. They were also more inclined to believe the system can work and to see it as less unfair than their civilian counterparts.

d. There were some differences among the divisions. The division with the highest job satisfaction cluster mean (4.72) had the lowest mean score on general reaction to TPMS (3.19) and fairness (2.49), and among the lowest mean score on usefulness as a management tool (3.27). This same division scored high on understanding (4.16). This can be contrasted with the division having the lowest job satisfaction mean (3.89), which had much higher scores for general reaction (3.63), fairness (3.70), and usefulness as a management tool (3.87). This division did not have as high a mean on understanding as the high satisfaction division (3.85 versus 4.16). The division with the higher job satisfaction scores seems to be more contented with the status quo and less accepting of TPMS (which is a change) than the division with not as high a level of job satisfaction. The latter is more accepting of TPMS and more confident TPMS will be a positive change, possibly reflecting a hope of better things to come.

3.4 RESPONSE TO FIRST OBJECTIVE

3.4.1 Objective 1: To determine how the TRASANA Technical Staff (TTS) perceives TPMS.

a. Study Question 1: What is the TTS general attitude toward TPMS?

The overall reaction to TPMS was mixed. There was considerable concern and uncertainty. One concern was that some staff members manipulate the system rather successfully and thereby achieve higher productivity indexes (PI) than others who are just as productive, but do not "play the system" as well. There was rather deep concern about how the results will be used - especially the individual PI - and the resulting effect on morale and the spirit of cooperation so highly valued at TRASANA. Another concern was that TPMS may result in a long term decline in the willingness to take on high risk projects and to try to develop new methodologies. On the positive side, the detailed planning for all projects was seen as very positive, the majority think resources can be realistically estimated with more experience, and one-third reported that they had seen some positive effects of TPMS during its first year.

b. Study Question 2: What is the perceived motivational impact of TPMS?

Over half of the staff do not think TPMS will motivate people to be more productive, but they are evenly divided on whether TPMS will lead to a long term productivity increase. These results reflect the reality that TPMS will not have the same impact on all people. What the total impact will be depends to some extent on how the system is changed to maximize the positive and minimize the negative factors.

c. Study Question 3: Is TPMS seen as a viable managerial tool?

Yes. This was seen as the most positive aspect of TPMS. In particular, the detailed planning required for all projects in the form of the project plan was seen as very useful.

3.4.2 Organizational Climate

The TRASANA staff feel very positive about the organization, see their work as challenging, and like the high standards of TRASANA. There is a concern felt by both supervisors (division and branch) and staff that TPMS may have a detrimental effect on the spirit of cooperation and on the willingness of the staff to take on the very difficult and risky projects that in the past have been welcomed as a challenge.

CHAPTER 4

FINDINGS: FIRST YEAR RESULTS

This chapter presents the data collected to achieve the second study objective: to determine the degree to which TPMS is a meaningful measure of productivity. The following study questions for this objective are answered:

- To what extent do TPMS ratings agree with confidential supervisory productivity assessments for individuals and branches?

- Does TPMS have generally the same impact on the different TRASANA organizational divisions?

- What are the functional relationships among the scoring algorithm variables?

- What is the operational impact of the Productivity Measurement Board (PMB) and Product Review Board (PRB)?

The first step in this effort was to capture all of the first year scoring results. These data are presented for the sixty-one scored projects, as well as by TPMS scoring factor. The productivity index (PI) for individuals and organizational units (branches, divisions) is examined, the individual PI evaluation analysis is presented, and scoring problems revealed from analysis of the first year scoring data are discussed. The analysis of the TPMS functioning components addresses the Product Review Board and Productivity Measurement Board. The results of the mathematical analysis of the TPMS scoring algorithm are in this chapter. Since the mathematical analysis is somewhat technical and requires basic familiarity with mathematics and error analysis techniques, it is in appendix E.

4.1 FIRST YEAR SCORING RESULTS

4.1.1 Scoring Results by Project

Table 4-1 shows the project scores (PS) for the 61 projects which were completed and scored between 1 October 1980 and 30 June 1981. The highest was 2.376 and the lowest was 1.046. The mean project score was 1.681 and the standard deviation was 0.258. The maximum possible PS is 2.614, and the minimum is 0.504. The baseline score for a project on time, on resources, and of average quality (defined as 1.3) is 1.56. The score assigned to the projects on each scoring variable and the productivity points (PP) earned also are shown in the table.

4.1.2 Manpower Scoring Results

The manpower utilization factor (R) is scaled as a piecewise linear increasing function of the ratio of planned (R_P) to actual (R_A) manpower for the project. It has 0.7 as the lower limit and 1.2 as the upper limit value. A detailed discussion of the scoring algorithm is in appendix E.

TABLE 4-1

FIRST YEAR SCORING RESULTS

PROJECT	PS	R_P	R_A	R	C	T	Q	PP
1	1.046	6.50	9.12	0.879	1.0	0.8	1.24	9.54
2	1.184	12.00	9.83	1.072	1.0	0.8	1.15	11.64
3	1.236	3.60	3.33	1.030	1.0	0.8	1.25	4.12
4	1.271	9.00	13.16	0.861	1.0	1.0	1.23	16.73
5	1.329	1.75	2.43	0.883	1.0	0.943	1.33	3.23
6	1.370	18.00	18.46	0.980	0.9	1.0	1.30	25.40
7	1.370	2.00	23.07	0.990	1.0	1.0	1.15	31.60
8	1.404	3.60	12.02	0.900	1.0	1.0	1.30	16.88
9	1.417	14.00	11.02	1.085	1.0	0.8	1.36	15.61
10	1.431	2.55	16.84	0.980	1.0	0.9	1.35	24.09
11	1.454	8.20	10.85	0.904	1.0	1.0	1.34	15.77
12	1.457	8.50	6.71	1.084	1.0	0.8	1.40	9.78
13	1.482	11.50	14.76	0.915	1.0	1.0	1.35	21.88
14	1.495	16.77	12.33	1.106	1.0	0.867	1.30	18.44
15	1.525	74.00	15.70	1.086	1.0	1.0	1.17	23.94
16	1.532	35.50	13.63	1.200	1.0	0.8	1.33	20.88
17	1.545	19.25	20.04	0.990	1.0	1.0	1.30	30.96
18	1.554	92.00	97.92	0.981	1.0	1.04	1.27	152.20
19	1.560	74.00	69.12	1.040	1.0	1.0	1.25	107.83
20	1.563	37.25	38.91	0.987	1.0	1.1	1.20	60.81
21	1.575	8.00	8.76	0.972	1.0	1.0	1.35	13.79
22	1.593	10.00	9.86	1.006	1.0	1.1	1.20	15.71
23	1.600	5.00	2.37	-	-	-	-	3.79
24	1.616	2.00	1.39	1.122	1.0	1.0	1.20	2.25
25	1.626	68.50	61.28	1.042	1.0	1.0	1.30	99.63
26	1.626	7.90	8.30	1.050	1.0	1.0	1.29	13.50
27	1.630	6.00	3.89	1.140	0.99	1.0	1.20	6.34
28	1.638	34.00	33.70	1.004	1.0	1.0	1.36	55.19
29	1.640	13.00	12.92	1.002	1.0	1.0	1.36	21.14
30	1.643	35.00	32.77	1.025	1.0	1.086	1.23	53.85
31	1.644	76.00	115.11	1.000	1.0	1.0	1.37	189.24
32	1.659	2.70	2.81	0.988	1.0	1.0	1.40	4.66
33	1.660	37.00	37.26	0.998	1.0	1.027	1.35	61.85
34	1.668	12.00	11.86	1.010	1.1	1.0	1.25	19.48
35	1.668	10.00	7.20	1.110	1.0	1.0	1.25	12.01
36	1.685	2.50	1.23	1.200	1.0	0.9	1.30	2.07
37	1.702	125.00	111.51	1.043	1.0	1.0	1.36	189.79
38	1.718	2.80	2.09	1.101	1.0	1.0	1.30	3.59
39	1.724	6.75	5.40	1.080	1.0	1.0	1.33	9.31
40	1.754	4.50	3.75	1.067	1.0	1.0	1.37	6.58
41	1.773	10.00	6.58	1.137	1.0	1.0	1.30	11.67
42	1.785	18.00	15.86	1.086	1.0	1.0	1.37	28.31
43	1.800	11.20	4.40	1.200	1.0	1.0	1.25	7.92
44	1.820	13.00	10.75	1.069	1.0	1.0	1.42	19.59
45	1.850	18.00	12.88	1.078	1.0	1.1	1.30	23.82
46	1.873	63.00	57.74	1.020	1.1	1.07	1.30	108.14

TABLE 4-1 (Continued)

PROJECT	PS	R_P	R_A	R	C	T	Q	PP
47	1.890	10.00	8.01	1.080	1.07	1.01	1.35	15.15
48	1.909	2.55	0.96	1.200	1.0	1.02	1.30	1.83
49	1.914	8.00	7.72	1.000	1.074	1.1	1.35	14.78
50	1.920	2.50	0.64	1.200	1.0	1.0	1.33	1.23
51	1.923	14.50	7.71	1.187	1.0	1.0	1.35	14.83
52	1.930	44.00	33.84	1.044	1.0	1.1	1.40	65.27
53	1.936	6.00	4.08	1.128	1.0	1.1	1.30	7.90
54	1.954	2.70	1.60	1.163	1.0	1.0	1.40	3.13
55	2.059	9.00	2.74	1.200	1.1	1.0	1.30	5.64
56	2.076	9.00	4.80	1.187	1.02	1.1	1.30	9.97
57	2.095	15.60	9.56	1.155	1.1	1.057	1.30	20.03
58	2.096	1.50	0.84	1.176	1.0	1.1	1.35	1.76
59	2.104	11.26	6.65	1.164	1.0	1.1	1.37	13.99
60	2.138	16.00	4.49	1.200	1.0	1.1	1.35	10.67
61	2.376	1.20	0.10	1.200	1.0	1.1	1.50	0.24
MEAN	1.681	19.76	17.95	1.063	1.008	0.997	1.309	29.36
ST DEV	0.258	25.43	25.36	0.093	0.031	0.083	0.069	41.87

NOTE: Not all scores are shown to the same number of significant digits because there was variation in the officially recorded scores.

KEY: PS-Project Score; R_P-Planned Manpower; R_A-Actual Manpower; R-Manpower Utilization; C-Computer Utilization; T-Timeliness; Q-Quality; PP- Productivity Points

a. The R factor results for 60 projects are shown in histogram form in figure 4-1.[1] The planned manpower ranged from 1.2 to 125 technical man months (TMM) with an average of 19.8. The actual manpower utilized varies from 0.1 to 115.11 TMM with an average of 17.9. Values for R ranged from 0.861 to the maximum allowable score of 1.2. The average was 1.063 and the standard deviation was 0.093. Twenty-five percent of the projects received an R score of approximately 1.0, which means the R_A was equal to the R_P. Two thirds of the projects had an R value greater than 1.0, which means they received a

---

[1] *As shown in table 4-1, one project did not have scores on R, T, Q, or C.*

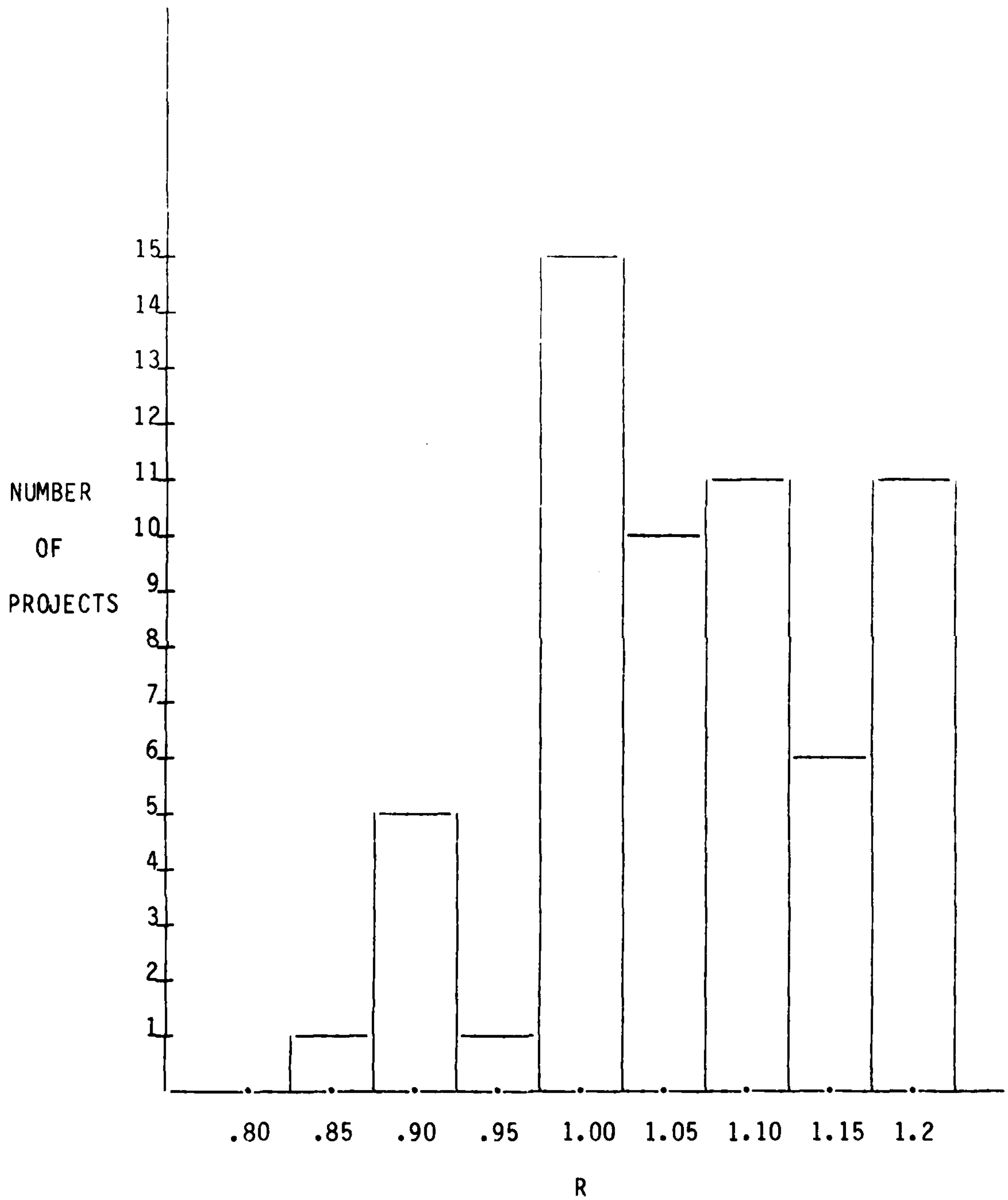


Figure 4-1. First Year Scores for Manpower Efficiency (R)

"bonus" for R_A being less than R_P. Thirteen percent of the projects received the maximum R score of 1.2. To achieve the maximum R value, R_A must be less than or equal to one half of R_P. For the eight projects achieving the maximum R score, R_P varies from 2.033 to 12.0 times R_A.

b. If manpower estimates were reasonably accurate, the R distribution would be symmetrical around a mean of approximately 1.0, or most projects would have R_A approximately equal to R_P, with as many projects falling under as over the estimates. The probability that 43 of 60 projects (72 percent) would be under the estimated manpower requirement, assuming accurate estimations, is approximately 0.0004, or 4 times in 10,000. It can only be concluded that R_P was consistently overestimated.

c. There were several factors driving the consistent tendency to estimate required manpower high. Such accurate estimating as required by TPMS was new to the staff, and it takes some experience to get good at such a difficult task. Even with practice, estimates for complex, risky projects can never be completely accurate. When this is considered along with the fact that TPMS rewards overestimating with higher project scores, and therefore tends to punish either accurate or underestimating with lower scores (less reward), the driving force for overestimating becomes clear. The only explicit punishment in this function is for substantially underestimating manpower requirements which results in earning points below the base level of 1.0.

d. There are two system checks on R_P. Before a project plan can be entered into TPMS for scoring, the responsible division chief must show that his total manpower committed for all approved projects is not more than 125 percent of available manpower (125 percent rule). The other check is that the project plan must be reviewed and approved by the PMB. The purpose of this review is to check for realistic and accurate resource estimates and schedules. These two checks were not effective in controlling the manpower estimates.

4.1.3 Product Timeliness Scoring Results

When a project plan is submitted to the PMB for approval, it must contain a proposed schedule for scoring. When the project is completed (including either a Q score of 1.3 or higher or PRB approval for release), a score is assigned based on the relationship of the actual completion date to the planned completion date. The values for this variable are 0.8 to 1.1, with 0.8 being the lowest score for being late and 1.1 being the highest possible score for completion ahead of schedule.

a. The distribution of the timeliness (T) scores is shown in figure 4-2. Thirty-two of the 60 projects (53 percent) received a score of 1.0, which means they were delivered exactly on schedule. Six (10 percent) received the minimum score of 0.8 and 11 (18 percent) received the maximum value of 1.1. The mean T was 0.997 with a standard deviation of 0.083. Ten (17 percent) were late, and 18 (30 percent) were delivered ahead of schedule.

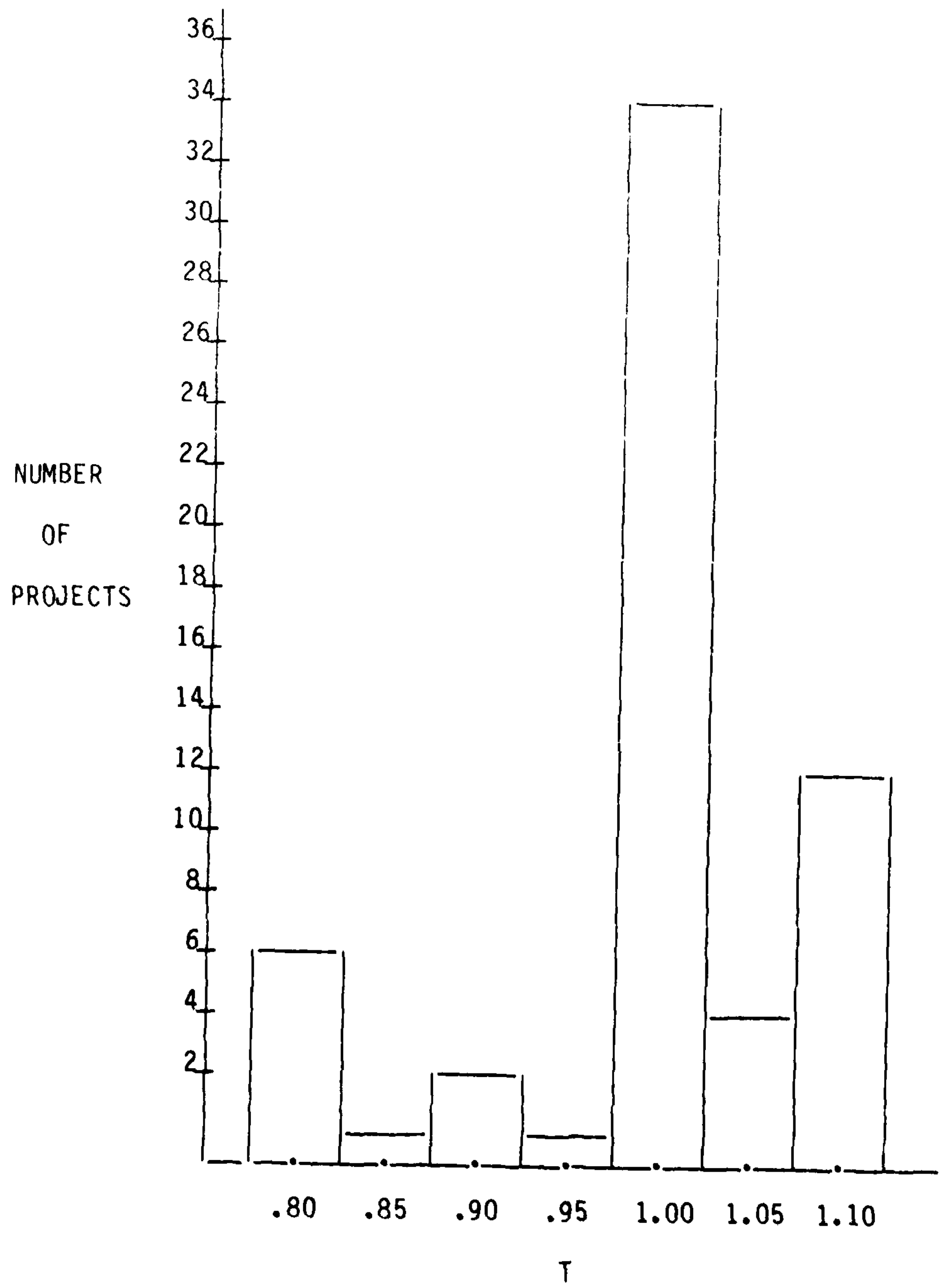


Figure 4-2. First Year Scores for Timeliness (T)

b. One reason for the large number of projects being delivered on schedule is external deadlines. In many cases the reports are of limited value if not available to decision makers by a specified date, even if the scope of the analysis must be reduced. In such cases, the project scope must be tailored to the deadline and the report published on time.

4.1.4 Product Quality Scoring Results

When a project is completed, the Product Review Board (PRB) convenes to review the product generated by the project. This product is typically a report which is distributed to the board members in advance. The three permanent board members assign a product quality score (Q) that can range from 1.0 to 1.5. A minimum score of 1.3 is required for the report to be released for distribution outside the activity. Reports achieving a Q score of less than 1.3 require revision before release. The Q score for the product does not change as a result of PRB required revisions.

a. The histogram in figure 4-3 shows the distribution of the 60 quality scores for the first year. The distribution is negatively skewed with a mean of 1.309 and a standard deviation of 0.069. The scores ranged from 1.15 to 1.5. Seventeen scores (28 percent) were below the established minimum score for release (1.3). This means that the PRB required substantial revisions or "polishing" on over one-fourth of the products. This quality control function of the PRB thus had a major impact during this nine month period.

b. The PRB was not new with TPMS, but assigning numerical scores was. It has long been recognized as a need because of the highly technical nature of the analysis performed by TRASANA which must typically be documented in a report understandable by readers who often do not have the training necessary to interpret a very technical report. This requirement to present a large amount of data and complex analyses in relatively non-technical terms accounts for many of the scores below 1.3.[2]

c. The quality score is the only variable in the TPMS scoring function which is totally independent of the project team. While the team is present during the review and actively interacts with the board, the board decides the score. The fact that Q is independent of the team would seem to make this factor a candidate for heavier weighting in the PS; however, there are three formidable problems with Q: (1) adverse morale impact, (2) historical difficulty in achieving high reliability for such scoring tasks, (3) the inability of a few individuals (any individuals) to be thoroughly knowledgeable of all the relevant issues and problems involved in conducting the large number of diverse projects typical for a year at TRASANA.

---

[2] *The person who served as PRB board chairperson during the first year was interviewed upon his departure from TRASANA shortly after the first scoring period ended. It was his conclusion that this was the case.*

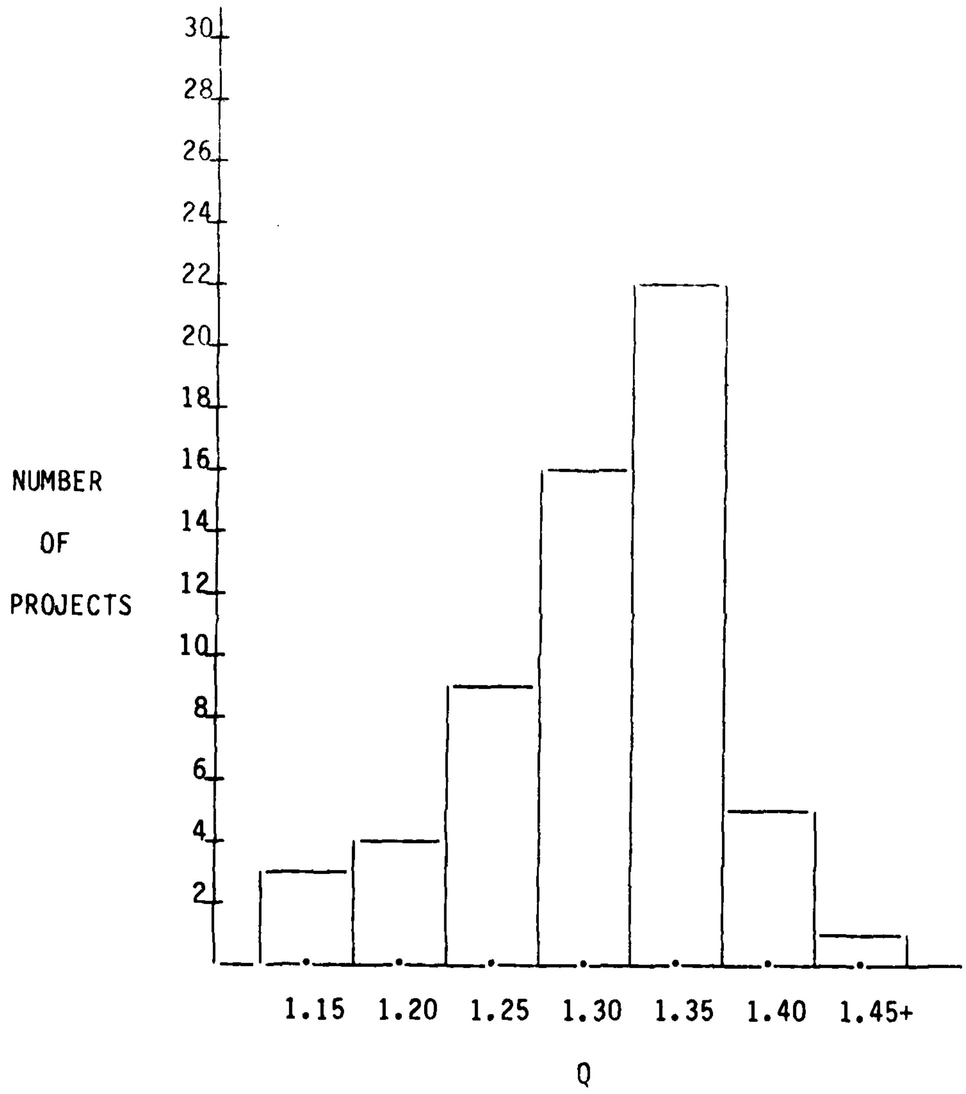


Figure 4-3. First Year Scores for Quality (Q)

(1) The most serious of the problems with Q is its adverse impact on morale. This problem was consistently evident from division chiefs to lower grade team participants. In confidential interviews with division chiefs, branch chiefs, and study managers, the comments about the PRB and Q score were like the same recording replayed. The criticisms were that the board in general does not have the broad range of technical expertise and historical knowledge of each project to comprehensively evaluate the project effort. Furthermore, there is no consistent PRB assessment and consideration of innovation, creativity, risk, and potential impact of each product. The bottom line is that the TTS perceives that products are scored based on format, composition, and editorial considerations, not quality of analytical effort.

(2) There was no way for the evaluation team to objectively determine the exact process used to arrive at a Q score for a product; however, available research on grading in general clearly documents the fact that there is considerable error in making fine discriminations of the type required of the PRB.[3] While the magnitude of error can be reduced by the use of detailed grading criteria or standards, it can not be eliminated. The absence of such standards in scoring by the PRB lends credence to the general contention that there is arbitrariness in scoring.[4]

(3) The basic issue reduces to whether the three permanent voting PRB members augmented by *ad hoc* members for specific projects can consistently apply the same exact standards with the same level of objectivity and accuracy to each and every product. If this can in fact be done, the Q score accurately reflects the relative product quality. If not, adjustments must be made in the Q scoring procedure. The fact that there is very little confidence in the relative Q ranking of products by those who scored high as well

---

[3]*In general, it has been shown that professionals who grade or score works, and then repeat the task with the same products after a substantial time lapse do not show a high level of agreement with themselves. Furthermore, papers scored by different professionals do not receive the same grades. Factors such as knowledge of author also affects grades, regardless of quality. A discussion of these and other related grading problems can be found in many educational psychology books such as Gage and Berliner (1975), or introductory measurement and evaluation books such as Mehrens and Lehmann (1973).*

[4]*TRASANA memos 5-4 (Dec 80) and 5-5 (May 81) do in fact provide policy guidance and standards for product scoring. Unfortunately, requiring the PRB to determine if appropriate methodology was employed does not provide any help in making that determination. The predominance of scientific evidence suggests that the PRB (or any such body) is not capable of making such fine distinctions as the difference between a 1.32, a 1.33, and a 1.34 report with a high level of consistency. This does not reflect negatively on the PRB members, but merely points out that they are attempting to do a nearly impossible task.*

as low, and by nearly all of the supervisors in the organization, makes a strong case for revised procedures, even if only for employee morale. This is even more evident when it is realized that the only project to score 1.5 was not an analytical effort, and produced neither a report nor a briefing to be scored.[5]

4.1.5 Computer Efficiency Scoring Results

The computer efficiency factor (C) score is computed in a manner similar to the R score. The values for C range from 0.9 for underestimating computer requirements to 1.1 as the maximum efficiency score. Of the 60 projects receiving a C factor score, 51 scored 1.0, two were below 1.0 and seven were above 1.0. One project received the minimum score of 0.9, and four the maximum score of 1.1. This factor was relatively unimportant in the scoring because the unit of measure changed during the scoring period from central processing unit (CPU) to standard unit of processing (SUP). Since most project plans estimated computer requirements in CPU's, and there is no direct CPU-SUP conversion, the majority of projects were assigned the C score of 1.0 administratively.

4.1.6 Priority Factor Scoring Results

The priority factor (P) was included in the scoring formula as an incentive to strive for products meeting the needs of customers external to TRASANA. A value of 1.2 is assigned to all projects conducted in support of external customers, and all internal projects are assigned a priority that can vary from 1.0 to 1.2. This factor did not work well. The PMB was unable to find an equitable way to assign priorities. The P score thus became a constant value equal to 1.2 for all projects.

4.1.7 Project Score Results

a. The distribution of the 61 project scores for the first year of TPMS is shown in figure 4-4. The minimum possible PS is 0.504 and the maximum is 2.614. The negatively skewed distribution shows that projects tended to exceed the base score value of 1.56 (1.3 Q score and on planned resources and schedule). The mean score was 1.681 and the standard deviation was 0.258. One project received the base score, and 42 (69 percent) projects scored above the base score.

b. Table 4-2 shows the relative effect on the PS of the same unit change (0.1) in the Q, T, R, and C factors. The effect of the increase is less for the Q factor than for the other factors. This is because the Q scale has a greater range than the other factors, so a 0.1 change has less impact than it does on the more compressed range values. The proper strategy for maximizing PS is not to concentrate on Q, but to concentrate on T, R, and C.

---

[5] *This project was monitoring of the installation of smoke detectors in TRASANA. The next highest score was 1.42.*

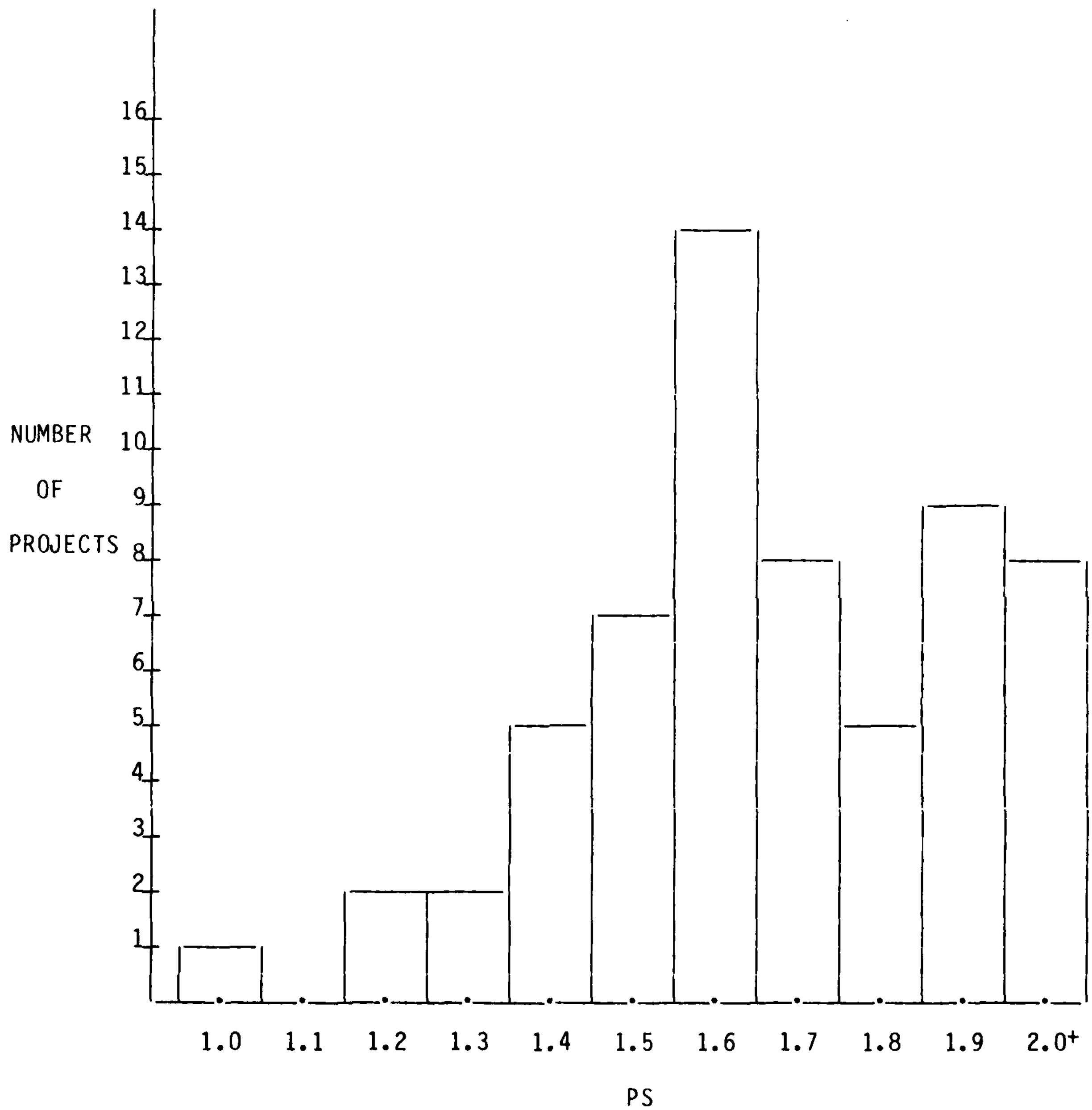


Figure 4-4. First Year Project Scores (PS)

The payoff for being early and under on manpower and computer usage is not only greater but easier to achieve since the team establishes the estimates. Note that a Q score of 1.4 or higher was achieved by only six projects (10 percent).

TABLE 4-2

SENSITIVITY OF PROJECT SCORE TO CHANGES IN FACTOR SCORES

CASE	P	Q	T	R	C	PS
BASE	1.2	1.3	1	1	1	1.560
Δ QUALITY	1.2	1.4	1	1	1	1.680
BEST TIME	1.2	1.3	1.1	1	1	1.716
$R_p = 1.5\ R_A$	1.2	1.3	1	1.1	1	1.716
$C_p = 2C_A$	1.2	1.3	1	1	1.1	1.716

4.1.8 Scoring Results by Division

Table 4-3 shows the mean and standard deviation for R, C, T, Q, and PS by TRASANA Division. The number of projects listed for each division includes scored projects for which the division was primarily responsible. For R, C, and T, the closer the mean is to 1.0, the more accuracy there was in the estimates. Futhermore, the more inaccurate they are on the positive side (overestimates), the higher the division PI. The smaller the standard deviation, the less spread there was in the scores for that factor. The average project scores are not the same as the division PI's. The PI is a weighted average and also includes overhead.

4.2 PRODUCTIVITY INDEX

The PS is designed to reflect the efficiency and effectiveness of a team of analysts in completing a specific project. Another way of stating this is that the PS is a productivity index (PI) for one project. The PS can be thought of as a team PI. TPMS defines PI as a weighted average of project scores. This weighted average is calculated for individuals (the average of productivity from different projects) and for organizational units. In the following discussion, the PI's calculated for the first year are summarized, and then the PI evaluation analysis is presented.

TABLE 4-3
AVERAGE PROJECT AND FACTOR SCORE BY DIVISION

DIVISION	PROJECTS	R	C	T	O	PS
A	6	1.019 (.093)	1 (0)	.983 (.098)	1.290 (.091)	1.571 (.353)
B	10	1.078 (.062)	1 (0)	1.029 (.046)	1.303 (.068)	1.736 (.185)
C	10	1.078 (.088)	1.029 (.044)	1.009 (.101)	1.292 (.044)	1.742 (.290)
D	8	1.049 (.125)	.986 (.035)	.995 (.095)	1.287 (.056)	1.583 (.210)
E	12	1.090 (.103)	1.008 (.029)	.975 (.097)	1.353 (.066)	1.744 (.294)
F	2	1.090 (.156)	1 (0)	.950 (.071)	1.30 (.071)	1.615 (.261)
G	3	1.071 (.072)	1.050 (.071)	1.035 (.049)	1.25 (.071)	1.696 (.153)
H	9	1.015 (.077)	1.008 (.023)	.99 (.078)	1.317 (.085)	1.629 (.260)
TOTAL	60	1.063 (.093)	1.008 (.031)	.997 (.083)	1.309 (.069)	1.681 (.259)

NOTE: The top number in the table is the mean value, and the number in parentheses is the standard deviation.

4.2.1 Individual PI Results

a. Figure 4-5 shows the distribution of individual PI's for the first scoring period. Only PI's including at least one man-month are included. This was necessary to eliminate spuriously high PI's that were primarily an artifact of system start up. For example, one individual received 0.11 productivity points on 0.07 TMM, and also collected 2.0 discretionary points (DP). The resulting PI of 30.14 was unrealistically inflated because the individual only charged a fraction of a month to scored projects. Inflated and unrepresentative PI's emerged in several instances because projects initiated under TPMS were not completed by the end of the first scoring period. This situation should occur less frequently in future years, since all projects will be covered by TPMS, and therefore the number of people not working on scored projects in any year should decline. There were 152 persons with PI's meeting the one TMM charged to scored projects criterion. The mean of the distribution was 1.689 and the standard deviation was 0.272. The values ranged from 1.09 to 3.49.

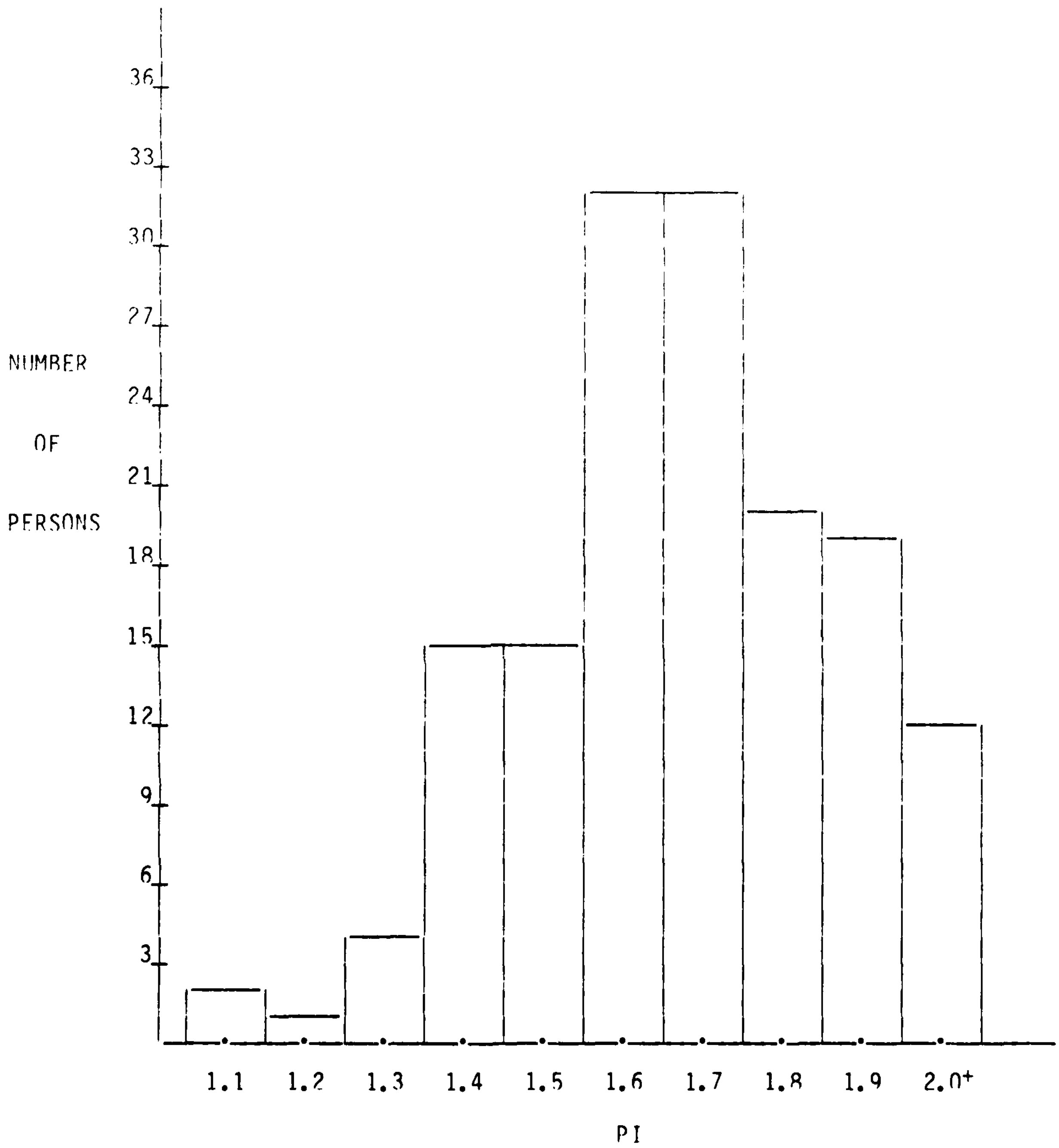


Figure 4-5. First Year Individual PI's (TMM $\geq$ 1.0)

b. A better understanding of the individual PI can be achieved by examining various groupings of individuals. Table 4-4 shows scoring information for the nine individuals with PI's exceeding 2.0. The project column in the table shows the number of scored projects to which the individual charged time, and TMM is total time charged against those projects. Two of the individuals achieved the high PI erroneously, illustrating the difficulty in monitoring the tedious and numerous numerical operations required by the system. Such errors were not atypical of the record keeping and computations.

c. PI's for the top nine scorers were computed without judgmental and discretionary awards, and are shown in the last column of table 4-4. These figures show that none of the nine would have achieved PI's over 2.0 without the additional bonuses and several approach the mean PI of 1.689. Since the bonus awards are at the discretion of individual supervisors who make the awards without specified performance standards, it is not clear what significance should be attributed to being a top scorer.

d. The top nine individual PI's were contrasted with the top ten scoring projects. Table 4-5 shows the results. The rows of the table refer to individuals (the top nine) and the columns to projects (the top ten). The entry in the table gives the percent of the individual's TMM which was charged to each specific project and the discretionary points which were awarded on each project. For example, the table shows that the person with the seventh highest PI (2.049) charged 85 percent of his TMM to the tenth highest scoring project. The interesting thing about the table is the large number of blank squares. For the most part, the highest scoring individuals did not work on the highest scoring projects.

e. In constrast to the top nine scoring individuals the nine lowest scoring individuals were also examined. Table 4-6 shows the scoring of these individuals. The three lowest scores were achieved by individuals who had left the activity by the end of the scoring period. Each of the three received negative judgmental awards. For all but one project, the lowest possible award was given, which was 30 percent less than the charged manpower award. Of the six still at TRASANA, the grade distribution is noteworthy: one GS-12, two GS-11's, and three GS-9's. The three GS-9's were interns (trainees). Recall that of the seven civilians in the top nine scorers, there was one GS-12 and six GS-13's. The relationship of grade and PI is examined more thoroughly below. To conclude the discussion of the bottom nine it should be noted that individual 7 achieved the low PI in an unusual way. After the productivity points had been computed and distributed, this intern continued to charge time to the project. This resulted in the PI decreasing from a rather respectable 1.65 to the 1.34 shown in the table.

f. In theory TPMS makes no distinction between military and civilian, the grade of the individual, or the division to which the individual belongs. The first year scoring data were analyzed to determine if there were in fact any such relationships.

TABLE 4-4. SCORING DATA FOR THE NINE HIGHEST INDIVIDUAL PI'S

INDIVIDUAL	RANK	NUMBER OF PROJECTS	TMM (SCORED PROJECTS)	PP	JUDGMENT AWARDS[c]	DISC AWARDS	DPP	PI	PI W/O JUDGMENT & DISC AWARDS
1	GS-13	3	1.39	4.850	0	2	2.5	3.490	1.690
2	GS-13	8	3.12	8.950	1	1	2.0	2.869	1.830
3[a]	GS-13	2	3.43	8.286	0	1	2.0	2.416	1.800
4	GS-12	2	11.10	24.850	2	0	0	2.240	1.656
5	GS-13	5	6.18	13.430	2	0	0	2.170	1.631
6	GS-13	1	24.11	50.530	1	0	0	2.096	1.644
7[b]	CPT	3	9.23	18.914	0	2	0.78	2.049	1.877
8	GS-13	3	7.27	14.877	1	2	0.825	2.046	1.805
9	CPT	1	5.13	10.380	1	0	0	2.023	1.702

[a]*DPP awarded for "exceptional contribution" on a project to which TMM was charged and for which team member PP's were awarded based on manpower charges. This violates TRASANA Memo 5-4 rules for awarding DPP.*

[b]*PI for this individual contains an arithmetical error. PP should 18.104 for a PI of 1.96. This person also received DPP on a project to which manpower was charged, but received no manpower PP for the project. This is also a violation of TPMS rules.*

[c]*All judgmental awards to these individuals were positive, which means they received more PP's than the award for their charged manpower. For this to occur, the equal number of PP's were withheld from other persons via negative judgment awards.*

TABLE 4-5. CONTRIBUTION OF HIGHEST SCORING INDIVIDUALS TO HIGHEST SCORING PROJECTS

PI (HIGHEST NINE) \ PS (HIGHEST TEN)	2.376	2.138	2.104	2.096	2.095	2.076	2.059	1.954	1.936	1.930
3.490						0.5 DPP				
2.869										
2.416										
2.240										
2.170					0.16%					
2.096										
2.049[a]										85%
2.046										68%
2.023										

[a]*This is one of the individuals whose PI contained inappropriate discretionary productivity points as well as a computational error.*

TABLE 4-6. SCORING DATA FOR THE NINE LOWEST INDIVIDUAL PI'S

INDIVIDUAL	RANK	NUMBER OF PROJECTS	TMM (SCORED PROJECTS)	PP	JUDGMENT AWARDS[b]	DISC AWARDS	DPP	PI	PI W/O JUDGMENT & DISC AWARDS
1	SP 5	1	4.45	4.860	- 1	0	0	1.092	1.560
2	GS-11	2	5.53	6.050	- 2	0	0	1.094	1.560
3	CPT	3	10.78	12.680	- 2	0	0	1.176	1.513
4	GS-11	2	3.93	5.071	0	0	0	1.290	1.290
5	GS-9	2	2.13	2.812	0	0	0	1.320	1.320
6	GS-9	3	7.07	9.394	0	0	0	1.329	1.329
7[a]	GS-9	3	2.54	3.395	0	0	0	1.340	1.340
8	GS-12	1	5.62	7.730	0	0	0	1.380	1.380
9	GS-11	1	1.14	1.570	0	0	0	1.380	1.380

[a] *This person continued to charge time to a project after it was scored. This increased his TMM and lowered his PI.*

[b] *All judgmental awards to these individuals were negative, and in 4 of the 5 cases they received the lowest award they before could be given, which is 30% below the award for their charged manpower. All three departed the activity before scores were computed.*

(1) Table 4-7 shows the comparison data for military and civilian PI's. There was no difference in the treatment of military and civilians by TPMS.[6]

TABLE 4-7. COMPARISON OF MILITARY AND CIVILIAN PI

CATEGORY	N	MEAN PI	STANDARD DEVIATION	MINIMUM	MAXIMUM
MILITARY	29	1.66	0.22	1.092	2.096
CIVILIAN	123	1.70	0.28	1.094	3.49

(2) The PI summary data for three civilian grade categories are in table 4-8. These data show a relationship between grade level and individual PI. There was a statistically significant difference between the GS-7 to 11 group and the GS-13 group.[7] Either the PI is measuring a real difference in productivity by grade level, or there is a difference in the treatment of grade levels in the application of TPMS.

TABLE 4-8. COMPARISON OF PI BY CIVILIAN GRADE

GS LEVEL	N	MEAN PI	STANDARD DEVIATION	MINIMUM	MAXIMUM
7 - 11	25	1.57	.22	1.09	1.98
12	43	1.68	.19	1.39	2.24
13	49	1.77	.39	1.40	3.49

---

[6] *This comparison was made by t-test; $t = 0.62$, $p = .536$.*

[7] *The statistical test applied was Analysis of Variance (ANOVA); $F = 4.232$, $p < .02$. The multiple range test applied was Scheffe at the .05 level.*

(3) Table 4-9 gives the individual PI comparisons by TRASANA division. GS-14's and 15's are not included because their PI's are primarily branch or division PI's. Statistical analysis revealed no significant differences among division mean PI's.[8] There was no advantage as far as individual PI goes to being in any one of the eight divisions.

TABLE 4-9. COMPARISON OF PI BY DIVISION

(Civilian - GS-13 or Lower)

DIVISION	N	MEAN PI	STANDARD DEVIATION	MINIMUM	MAXIMUM
A	19	1.65	0.25	1.09	2.10
B	17	1.70	0.11	1.56	1.97
C	18	1.71	0.21	1.40	2.24
D	13	1.68	0.12	1.54	1.85
E	13	1.84	0.62	1.40	3.49
F	7	1.64	0.39	1.40	2.42
G	12	1.71	0.23	1.38	1.87
H	19	1.63	0.24	1.29	2.05
TOTAL	118	1.69	0.29	1.09	3.49

4.2.2 Individual PI Evaluation Analysis

The extent to which the individual PI reflects productivity can only be determined by comparing the PI results with an independent productivity measure or criterion measure (CM). The results of the PI evaluation analysis are presented next. Since CM data are sensitive and were collected under the agreement with branch chiefs that the data would remain confidential, the data presented are summarized sufficiently to avoid the possibility of individual or branch identification.

[8] *This comparison was by ANOVA, $F = 0.749$, $P = 0.63$.*

a. Table 4-10 shows the number of people utilized in the PI evaluation analysis by division. The 125 persons who had both a useable PI (TMM > 1.0) and CM score was 83 percent of the total number with a useable PI, which excludes division chiefs, branch chiefs, deputies, and technical advisors.[9]

TABLE 4-10. PI EVALUATION PARTICIPATION BY DIVISION

DIVISION	PERSONNEL AVAILABLE FOR ANALYSIS[a]	NUMBER WITH USEABLE PI[b]	NUMBER WITH USEABLE PI AND CM (%)
A	26	25	20 (80)
B	20	20	18 (90)
C	27	20	20 (100)
D[c]	34	21	13 (62)
E	32	21	17 (81)
F[d]	24	7	7 (100)
G	17	14	10 (71)
H	31	23	20 (87)
OVERALL	211	151	125 (83)

[a]*TTS excluding division chiefs, branch chiefs, deputies, technical advisors, and identified new arrivals.*

[b]*More than 1.0 TMM charged to scored projects.*

[c]*One branch had no branch chief.*

[d]*There were only two scored projects in this division.*

---

[9]*The PI sample was representative of the TTS. The representativeness of the PI sample by grade was tested by Chi-Square.* $X^2 = 4.08$, $df = 7$, *which is not statistically significant at the* $p < .05$ *level* ($X^2_{7,.05} = 14.07$). *Table C-11 in appendix C shows the composition of the PI sample and TTS by grade.*

b. CM data were collected from nineteen branch chiefs. The CM consisted of two separate data elements. First, the branch chief rank ordered the individuals in his/her branch on the basis of overall job performance. Next, each person in the branch was assigned a numerical rating on a scale of 1 to 10. The scale was defined as 10 being a "superstar," and one being the opposite of 10. Five was defined as the average for TRASANA. The rank order was obtained for computing a rank order correlation with the PI rank order. The rating was obtained to determine the level of agreement of branch chiefs in rating standards.

c. Table 4-11 shows the percentage of analysts rated in each of the 1 to 10 categories by the branch chiefs. The total number of ratings in the table is 169, which includes those individuals rated by branch chiefs who did not enter the analysis due to the lack of a useable PI. In all but three branches, CM measures were obtained before PI's were computed. The mean rating by branch ranged from a high of 8.39 to a low of 6.25, with a mean of 7.07, a standard deviation of 0.70 and a median of 7.0. These data show that the ten point confidential rating achieved the desired discrimination of performance levels.

d. A rank order correlation was computed between the CM rank order and the PI rank order for nineteen branches. Table 4-12 shows the resulting correlation coefficients from highest to lowest. The N is the number of individuals in each branch who had both of the necessary scores for the correlation. The third column in the table shows the necessary absolute value for the N for statistical significance at the .05 level. The small numbers involved in the individual correlations do not provide the stability desired for a high level of confidence in the results. The one clear point from these data is the tremendous variability. Only one shows a significant positive correlation, and it is offset by a significant negative correlation. These data fail to support the PI as a reliable indicator of individual productivity.

e. While the branch by branch rank order correlations of CM and PI reveal that there is no strong positive relationship, the fact that the number of pairs in each correlation is small may mask a low order trend. To test for this possibility, CM ratings (1 to 10) were used. Within each division, all CM ratings were listed from high to low and divided into three categories. The rationale was that these confidential supervisory ratings could be combined within a division and accurately identify three relative levels of performance - high, medium, and low. For each division the ratings were examined for cut points which would yield three approximately equal groups. This procedure resulted in a division by division categorization that varied somewhat among the divisions because the branch chief ratings were not a forced distribution.

f. After the three job performance categories were established for each division, the PI rank order for the division was used to form three categories with the same number of individuals, but using PI rank order instead of CM ratings. For example, the CM ratings for one division resulted in six individuals in the high performance category, 10 in the medium, and five in the low category. These 21 persons were then rank ordered from high

TABLE 4-11

BRANCH CHIEF CONFIDENTIAL RATINGS OF TRASANA ANALYSTS

RATING	N	%
10 (HIGHEST)	21	12.40
9	24	14.20
8	22	13.02
7	35	20.71
6	30	17.75
5	26	15.38
4	4	2.37
3	6	3.55
2	1	0.59
1 (LOWEST)	-	-
TOTAL	169	99.97[a]

[a]*Total does not equal 100% due to rounding.*

to low using their individual PI. The top six were put in the high PI category, the next 10 in the medium, and the last five in the low category. The two classifications, one by CM and one by PI, were then compared to see how different the two groups were. Figure 4-6 shows the percentage of individuals by division whose classification by PI rank order was different than by CM rating. The percentage misclassified by PI varies from 45 to 80 percent. The misclassifications did not exhibit any detectable pattern. Several individuals moved from one to the other extreme category.[10] These data show that the individual PI does not accurately categorize individuals into high, medium, and low performance groups.

---

[10]*The data for this analysis are not provided as agreed with branch chiefs. The PI categories showed marked variance from the CM categories, and the movement of individuals was not minor. In fact, 16.7% of the individuals went from the high to low group or vice versa, and a total of 58% changed categories.*

TABLE 4-12

CORRELATION OF PI AND CM RANK ORDER OF ANALYSTS BY BRANCH

CORRELATION COEFFICIENT	N	REQUIRED FOR SIGNIFICANCE AT .05 LEVEL	STATISTICAL SIGNIFICANCE
.90	9	0.683	< .01
.85	5	1.000	NS
.80	5	1.000	NS
.80	4	1.000	NS
.73	6	0.886	NS
.71	6	0.886	NS
.57	7	0.786	NS
.53	6	0.886	NS
.47	7	0.786	NS
.36	10	0.648	NS
.36	8	0.738	NS
.30	5	1.000	NS
.26	7	0.738	NS
.17	7	0.786	NS
.10	7	0.786	NS
-.02	10	0.648	NS
-.29	7	0.786	NS
-.71	6	0.886	NS
-1.0	3	1.000	.05

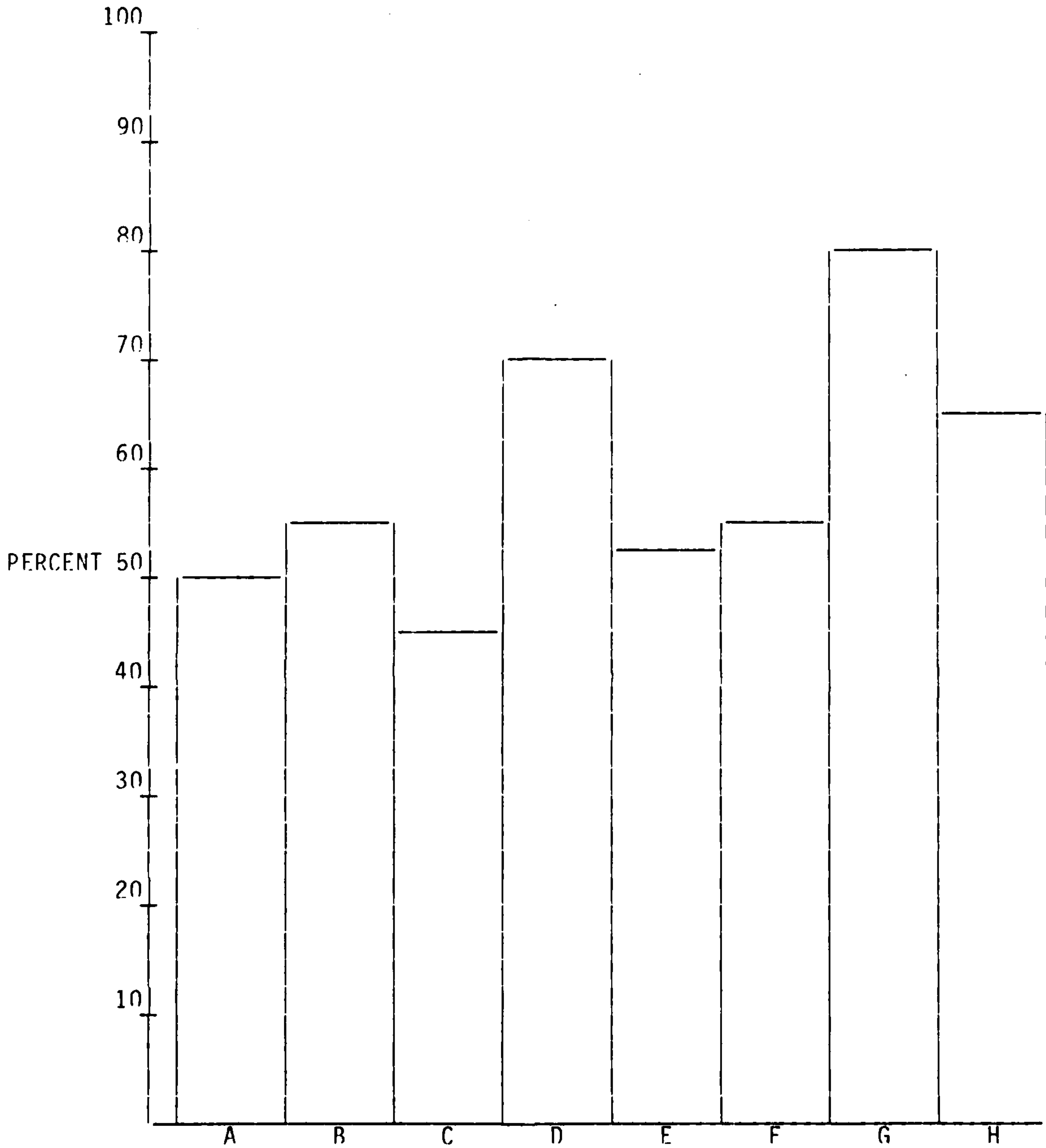


Figure 4-6. Individual Performance Misclassified by PI

4.2.3 Individual PI Difficulties

Conceptually, the PI for an individual is the PS modified by supervisory judgment to better reflect the individual's level of contribution to the PS. This adjustment takes the form of increasing or decreasing the PS by up to 30 percent for the individual.

a. One real difficulty with quantifying the individual contributions to the PS is that the PS actually reflects much more than individual effort and therefore so does the PI. The PS is essentially a team score. Many of the tasks in a project are actually a team effort which are virtually impossible to divide into individual contributions. Furthermore, the team performance is related to many things that any given individual on the team may have very little or no say so in, such as who is on the team, how the project is planned and executed, project difficulty (and other risk factors), and project leader decisions.

b. The adjustment factor itself is part of the problem. Plus or minus 30 percent of a low score may still be a low score, and the opposite in the case of a project scoring very well. A high performer on a poorly planned and executed project, or a poor performer on a very well managed project will have PI's which do not likely reflect the true difference in performances. This situation is further complicated by the fact that the supervisor is playing a zero-sum game. To give someone more, someone else gets less, because the total will always be the same. The only discretion is in distribution of the available points. There was a lot of variation in how this was handled. The inconsistency with which supervisors approached this situation was a major part of the PI problem. Some supervisors tried to make the PI's reflect their judgments, and others made awards almost entirely on manpower charges without applying the individual contribution adjustment factor.

c. The same inconsistency applied to the use of discretionary productivity points. The lack of understanding of the potency of DPP's was evident in some cases, which is why some PI's were unrealistically exaggerated. It has already been pointed out that the top nine scorers would not be top scorers without bonus awards. There is no problem *per se* in this fact. The problem is that the lack of consistency makes the winners and losers too dependent on factors other than performance.

d. A supervisor has at his disposal, the power to substantially alter an individual's PI by the application of the ± 30 percent rule, and the award of DPP, if he chooses to do so. The fact that supervisors were so inconsistent in this regard made the first year PI's meaningless. Furthermore, the PI does not capture all of the relevant aspects of job performance in all cases, and reflects factors over which the individual has no control. The individual's PI is therefore subject to considerable misinterpretation and abuse. Since it is not a good individual performance indicator, and there is a performance appraisal system designed to document job performance, the individual PI should be eliminated. Its retention weakens the productivity measurement system, but yet does not provide a useful performance appraisal device.

4.2.4 Branch PI's

a. Branch PI's ranged from 1.17 to 1.81 with a mean of 1.48 and a standard deviation of 0.15. The scoring data by branch are in table C-12 in appendix C. The branch with the lowest PI had less than one TMM charged to scored projects during the scoring period.

b. There was a wide range of charges to overhead (OH) among the branches. The lowest charge was 0.5 TMM and the highest overhead charge was 50.34 TMM. The effect of overhead charges on branch PI's is illustrated in table 4-13. The effect OH has on any given PI depends on several factors -- the rules used in charging or not charging overhead, the total branch TMM charged to scored projects, and the fact that some branches have a mission requiring overhead activities. The erratic overhead charging pattern and tendency to hide actual overhead in project manpower charges dilutes the usefulness of the PI for branch comparisons.

c. The inclusion of overhead in the PI has a substantial effect. The value of what a branch produces is not lowered by overhead, but the PI is. The reason for including overhead in the PI is to keep branches from charging time to overhead instead of to projects to affect the R score. The result is that overhead is not being accurately documented, and since R is not being controlled, overhead is being charged to projects. Overhead charging has to be controlled, but including it as a PI decrement makes the organizational PI less accurate.

4.2.5 Division PI's

Division PI's ranged from 1.32 to 1.60. Table 4-14 shows the division PI's with and without OH, and reveals the same difficulty of inconsistent OH charging as the branch figures. Division OH charges ranged from 11.10 to 108.16 TMM. More complete scoring data by division are in table C-13 in appendix C.

4.3 TPMS MANPOWER COSTS

There is no way to determine exactly how much manpower was expended on TPMS during the first scoring year. Some of the things required by TPMS are also required for good management, such as a detailed project plan. Such activities cannot legitimately be considered TPMS costs. Activities such as calculating PI's and distributing PP's are pure TPMS costs. A separate manpower charging code was established for documenting time spent for TPMS activities. In addition to the above mentioned calculations, time spent in PMB activities was also charged to TPMS. Table 4-15 shows the TMM charged to the TPMS overhead account during the individual scoring year as well as to PRB activities. As mentioned earlier, the PRB existed before TPMS. The PRB charges are shown separately because it represents a substantial investment and is an important part of TPMS, even though it is not a pure TPMS cost. It should be noted that most of the first year PI calculations were performed after the end of June, and therefore are not reflected in the figures in the table.

TABLE 4-13

BRANCH PI'S CALCULATED WITH AND WITHOUT OVERHEAD CHARGES

BRANCH	REGULAR OVERHEAD	PI WITH OVERHEAD	PI W/O OVERHEAD
1	4.06	1.81	1.86
2	6.75	1.64	1.69
3	3.02	1.63	1.72
4	0.50	1.63	1.72
5	7.80	1.62	1.70
6	2.71	1.59	1.61
7	2.68	1.58	1.60
8	7.35	1.58	1.68
9	9.44	1.57	1.65
10	5.70	1.55	1.66
11	16.63	1.52	1.71
12	16.72	1.50	1.79
13	13.72	1.47	1.75
14	9.01	1.44	1.58
15	33.07	1.42	1.72
16	4.03	1.42	1.92
17	4.55	1.41	1.46
18	3.56	1.38	1.45
19	44.14	1.36	1.78
20	11.44	1.33	1.53
21	50.34	1.30	2.38
22	12.70	1.27	1.39
23	13.12	1.23	1.40
24	4.96	1.17	1.83
MEAN	12.00	1.48	1.69
SD	12.86	0.15	0.20

NOTE: Branches are arranged from highest to lowest PI.

TABLE 4-14. DIVISION PI'S CALCULATED WITH AND WITHOUT OVERHEAD CHARGES

REGULAR OVERHEAD	PI WITH OVERHEAD	PI W/O OVERHEAD
11.46	1.60	1.63
26.55	1.57	1.66
27.39	1.56	1.76
11.10	1.55	1.59
66.74	1.46	1.67
78.67	1.39	1.81
108.16	1.35	1.70
31.57	1.32	1.60

NOTE: Divisions are arranged from highest to lowest PI.

TABLE 4-15. TPMS MANPOWER CHARGES FROM OCT 80 THROUGH JUN 81

DIVISION/ OFFICE	TMM FOR TPMS OVERHEAD	TMM FOR PRB OVERHEAD
A	0.84	1.16
B	1.60	1.74
C	2.80	1.72
D	11.62	1.42
E	3.38	0.60
F	6.80	2.90
G	0.69	1.50
H	1.54	1.50
DT	2.57	2.11
DA	0.35	2.99
TOTAL	32.19	17.64

4.4 FIRST YEAR SCORING PROBLEMS

There were numerous instances of computational inaccuracies and questionable scoring procedures in the first year scoring data. There were many reasons for these problems. The fact that it was a new system was certainly a large part of the problem, but by no means all of it. The record keeping and calculation requirements are cumbersome and had to be done manually. Such a situation invites errors. Furthermore, the diligent monitoring of so much data kept by several different offices would approach a full-time job. While some of the problems were due to system start-up, there is no reason to assume the second year data will be much improved unless considerable simplification, automation, and/or monitoring is undertaken. Selected examples illustrative of the kind of problems that were not uncommon are cited below.

4.4.1 Example 1 - Situation Not Covered by TPMS Guidance

This person received two discretionary awards consisting of 1.57 and 2.38 PP. No other TMM or PP were recorded for this individual. Since discretionary awards increase an existing PI, and this person had no PI, there was a situation where the routine application of the formula did not work. In this situation, a rather innovative solution was applied. The person's PI was computed as the average of the discretionary awards weighted by total project TMM, resulting in a PI of 1.78.[11]

4.4.2 Example 2 - Inaccurate Record Keeping/Computation

The project summarized in table 4-16 illustrates record keeping and computation difficulties. Values appearing in the first row of the table were taken from the TPMS project log (TMM and PP). The remaining data in the table come from the TPMS scoring sheets. While the two sources agree that the expended TMM was 23.07 by the responsible division, the number of productivity points earned is different from the two sources. The scoring sheets show that 1.04 more PP were awarded than the project log. Two branches charged time to this project which summed to 24.54 TMM. This is more than 1.0 TMM over the value for expended TMM in the project log. If the larger expenditure were in fact correct, the R value would drop from 0.99 to 0.96, and the project score would change from 1.37 to 1.32. Also, the sum of PP awarded at the branch level was 32.796, which is inconsistent with both the division score sheet and the project log.

4.4.3 Example 3 - System Manipulation

a. This example illustrates a TPMS design difficulty. The scoring of one branch with regard to this project is shown in table 4-17. Again the values in the first row come from the project log and the other data from the

---

[11] *If the PI had been calculated using the individual's regular OH it is likely that the PI would have been much higher. Suppose for illustrative purposes that the person had 1 man-month of regular OH. The PI computed with this OH is 4.95.*

TABLE 4-16. TPMS SCORING SHEET SUMMARY FOR PROJECT NUMBER 7

	TMM EXPENDED	PRODUCTIVITY POINTS	PI
Project	23.07	31.60	1.37
Division	23.07	32.64	1.41
Branch	24.54	32.796	
X	8.26	10.98	1.33
Y	16.28	21.816	1.34

TABLE 4-17. TPMS SCORING SUMMARY FOR PROJECT NUMBER 33

	TMM EXPENDED	PRODUCTIVITY POINTS	PI
Project	37.26	61.86	1.66
Division	37.17	61.70	1.66
Branch	24.63	46.45	1.89
Person A	10.80	24.24	2.24
Person B	4.21	9.45	2.24
Person C	0.06	0.14	2.33
Person Total	15.07	33.83	NA

scoring sheets. The branch received a judgmental award of 5.57 PP for a PI of 1.89 on this project. Score sheets for the branch show only three individuals charging to this project. These three individuals do not account for 9.56 of the man-months charged by the branch nor 12.62 of the productivity points awarded to the branch. The assumption is that these 9.56 man-months were charged by the branch chief who then took a minimum negative judgmental award (judgmental factor of 0.70) with absolutely no loss in the branch PI (which of course is the branch chief's PI). This negative judgmental award to the branch chief allowed positive judgmental awards to the three individuals in the branch. These awards put two of the three individuals into the top nine scorers in the activity (PI $\geq$ 2.00) even though the project score was 1.66, which is below the mean activity PS of 1.68.

b. The problem illustrated in this example is that when a branch chief charges to a project, the points which he earns may be used as bonus points for other individuals in the branch. They are of no use to the branch chief, and can thus be distributed to other team members. It is to the team's advantage then to have the branch chief charge time to projects. When this situation does occur, it creates inequity for individuals on projects to which the branch chief does not charge time, or on projects with large numbers of participants to share the bonuses. The same kind of manipulation occurred in instances where team members departed the organization before PP's were distributed.

4.5 PRODUCT REVIEW BOARD (PRB)

Data concerning the PRB functioning came from the first year scoring results and interviews with division chiefs, branch chiefs, and PRB members.

a. The results of the Q score analysis are presented in 4.1.4. Essentially, the Q score distribution was statistically a better distribution than those for the other factors. What is not revealed by the distribution is the accuracy of scores, or how well the Q score captures actual differences in project quality. This element was addressed in the interviews.

b. The result of the interviews was a unaminous lack of confidence in the Q scores. There was no difference in responses by those who had and had not fared well in their Q scores. It is generally believed in the organization that the PRB grades projects primarily on such factors as report format and editorial considerations. Not one supervisor was found who felt the Q score adequately discriminated between project quality. This lack of confidence in the score given by the PRB is a substantial detriment to morale. Several branch chiefs expressed concern that reports were being tailored more and more to the PRB and therefore less to the proponent. Another concern which frequently surfaced was that larger projects are penalized because it is much more difficult to organize and write a readable report for large complicated projects.

c. As was pointed out in 4.1.4, the PRB has a very difficult, if not impossible, task in trying to make fine discriminations among very diverse projects and products. The major problem is seen as the grade and the grading process, not the PRB. There was general agreement (incidently, supported by first year data) that the PRB is necessary, but its current power and influence in the TPMS scoring process is not well received. A more definitive balance of power is needed between the responsible division (or line management) and the PRB (staff function).

4.6 PRODUCTIVITY MEASUREMENT BOARD (PMB)

Data for the analysis of PMB functioning came from the first year scoring results, interviews with PMB members, and observing one PMB meeting. The PMB is responsible for reviewing all project plans for accuracy in estimated resources and schedules. The primary emphasis in this analysis was how the PMB members made these judgments, and how effective they were in controlling the input estimates.

a. First year scoring results reported in 4.1.2 document the fact that the resource estimates were not adequately controlled the first year. Interviews with PMB members (division chiefs) revealed that the general opinion of the PMB members was that they did not know how to judge the validity of estimates from divisions other than their own. In fact, considerable frustration was expressed in regard to being asked to make such judgments. As a result of this uncertainty, the tendency was to only raise questions in cases where blatant inaccuracies were suspected, which was rather infrequent.

b. In many respects, the PMB members are faced with the same dilemma as the PRB members, in that no matter how hard they try, it is not clear that their task is reasonably achievable. Considering that over 100 projects were entered into the system the first year, and that 60 to 90 reports are generated by TRASANA every year, the magnitude of the problem comes more clearly into focus. In addition to the number of projects, there is considerable variation in what the projects involve. There is just not an adequate amount of time for division chiefs to acquire and process detailed information about each and every project and keep abreast of their own division responsibilities.

4.7 TPMS ALGORITHM ANALYSIS RESULTS

a. Numerical error analysis revealed that the numerical error associated with both the project score and the PI is 0.027. The contributing factors are errors in the scoring factors and the computational methods. Depending on how the PI is used, this may be an unacceptably high numerical error. For example, if a PI is computed as 1.689 (first year mean) the true value could vary from as little as 1.662 to as much as 1.716. At the top end of the scale, a PI of 2.00 could have a true value varying from 1.973 to 2.027.

b. One of the stated reasons for using a multiplicative rather than an additive scheme for assembling the project scoring factors is that an additive scheme is compensatory while a multiplicative scheme is not. In fact, a multiplicative scheme is compensatory to a certain degree. This necessitates adopting a minimally acceptable value, sometimes called a critical hurdle, for Q (currently 1.3). The resulting scoring function is a combination of a compensatory and conjunctive (critical hurdle) scheme.

c. The use of distinct ratio efficiency factors (R and C) which are weighted differently is intended to reflect efficient resource use and also to encourage substituting manpower with less expensive computer usage. However, the different weighting distorts actual resource costs and the use of distinct factors could actually encourage replacing some computer usage by more expensive manpower to maximize the project score.

d. The hierarchical distribution of PP (first to divisions, then to branches, and finally to individuals) constrains the awarding process to the extent that supervisors may not be able to award PP equally for comparable work in some cases.

e. Branch and division PI's are weighted averages of PP (including overhead) over a scoring period. Scored projects can carry-over from one period to the next while overhead would not. This could result in organizations having equal productivity over two years, but their PI's for each year would be different.

4.8 RESPONSE TO SECOND OBJECTIVE

4.8.1 Objective 2: To determine the degree to which TPMS is a meaningful measure of productivity.

a. Study Question 1: To what extent do TPMS ratings agree with confidential supervisory assessments for individuals and branches?

(1) Individuals. There was not agreement between individual PI's and confidential supervisory assessments. Furthermore, the individual PI is to some extent subjective, reflects factors beyond the individual's control, and is subject to misinterpretation and abuse. The individual PI was plagued with so many difficulties that even with substantial modifications, its usefulness will be questionable.

(2) Branch. On the branch level, the evaluation was not based on supervisory evaluations. Branch PI's (as well as division PI's) were much more meaningful than individual PI's. The most useful aspect of the PI is for organizational units. If input estimates are better controlled (see 4.1.2), and the problems with the quality score (see 4.1.4) and overhead (see 4.2.4) are resolved, the organizational PI will be a solid productivity indicator, especially over a several year period.

b. Study Question 2: Does TPMS have generally the same impact on the different TRASANA organizational divisions?

No. The major problem here was the treatment of overhead. There were considerable inconsistencies in the way overhead was charged. As TPMS is currently structured, organizational units are punished (PI is lowered) for charging time to overhead. When PI's are calculated without overhead, the resulting branch (and division) rank order is virtually uncorrelated with the complete (with overhead) PI rank order. The inclusion of overhead, which earns productivity points at a lower rate than the worst scoring project, punishes those organizational units with an overhead mission, and dilutes the usefulness of the organizational PI.

c. Study Question 3: What are the functional relationships among the scoring algorithm variables?

(1) The most important driving factor behind first year scores was the manpower utilization factor (R). The priority (P) and computer utilization (C) factors, except for a very few C scores, were constants in the scoring equation. The timeliness factor (T) equalled 1.0 for over half of the projects. The quality (Q) score had a lesser effect on the project score than the R, C, and T factors because of its range, its scoring distribution, and the multiplicative scoring function.

(2) A conservative estimate of the numerical error associated with both the project score and the PI is 0.027. Depending on how the PI is used, this may be an unacceptably high numerical error.

(3) The use of distinct ratio efficiency factors (R and C) which are weighted differently could actually encourage replacing computer usage by more expensive manpower to maximize the project score.

(4) The hierarchical distribution of PP (first to divisions, then to branches, and finally to individuals) constrains the awarding process to the extent that supervisors may not be able to award PP equally for comparable work in some cases.

(5) Branch and division PI's are weighted averages of PP (including overhead) over a scoring period. Scored projects can carry-over from one period to the next while overhead would not. This could result in organizations having equal productivity over two years, but their PI's for each year would be different.

d. Study Question 4: What is the operational impact of the Productivity Measurement Board (PMB) and the Product Review Board (PRB)?

(1) PMB. The most important function of the PMB was controlling input estimates for R, C, and T. It was not effective at this, especially in the case of R. This does not reflect poorly on the PMB. This way of controlling estimates was a good idea that did not work.

(2) PRB. The PRB appears to have done a good job in the area of quality control (28 percent of the reports were rejected the first time around), but the actual assigning of "grades" was less successful. The problem was that the PRB tried to make very fine distinctions between projects based on reviewing reports in a relatively limited amount of time.

CHAPTER 5

SUMMARY AND CONCLUSIONS

This chapter summarizes the findings of the Phase I evaluation of TPMS relative to the original design attributes of the system, and presents the conclusions supported by the analysis. Summary responses to the objectives and study questions are at the end of chapters three and four.

5.1 DESIRED ATTRIBUTES OF TPMS

a. Table 5-1 shows the results of the first year evaluation relative to the original design attributes of TPMS (see chapter 1).

TABLE 5-1. FINDINGS RELATED TO TPMS DESIGN ATTRIBUTES

ATTRIBUTE	FINDINGS
Relevant	Yes; Measures productivity
Rational	Yes; Except individual PI and DPP
Fair	Individual: Input/output mismatch Organization: Overhead mission disadvantaged
Affordable	?
Flexible	Undetermined
Simple	No
Transparent	No
Durable	TBD

(1) TPMS was judged to be relevant to the organization, its products, and to the jobs that produce the products. The classical definition of productivity is value of output over value of input, and captures both efficiency and effectiveness. The R and C factors of TPMS capture efficiency, and the T and Q factors capture effectiveness.[1] The weakness in terms of meeting the classical definition is in controlling estimated manpower, which is the single most important factor in establishing project worth under the cost-plus-incentive concept.

---

[1] *The P factor captures neither efficiency nor effectiveness. In terms of the classical definition of productivity, the P factor is questionable. It was not an issue in this analysis because it reverted to a constant value for all projects.*

(2) The system was judged to be conceptually sound, and therefore rational, but some of the component elements did not work as well as others. Basically, the inconsistent application of judgmental and discretionary productivity points had an undesirable impact on individual PI's. Furthermore, so many of the tasks in projects are actually group tasks, or team tasks, that there is no objective way to measure the relative contributions of individuals.

(3) The most frequently stated complaint about TPMS was that it is unfair. There is some basis in fact for this complaint. On an individual level, there is considerable potential for an input effort/output reward mismatch. For example, two persons with equal capabilities and putting forth equal time and effort on two equally scoring projects (i.e. equal project scores) may not earn productivity points at the same rate. This could occur for reasons totally beyond the individual's control, such as one study manager using judgmental awar's to reward "super" performers, and the other study manager distributing all productivity points based solely on manpower charged (i.e., everybody on the team earns at the same rate). On an organizational level, those branches and divisions with an overhead mission are disadvantaged.

(4) No judgment is offered on the affordability of the system because that is a decision for management to make after all the facts are available.

(5) The flexibility of the system was not determined. One problem related to flexibility is the disadvantage to units with an overhead mission.

(6) The system is not simple. The PI calculations are cumbersome and complex, and the staff generally does not understand how the various factors and components interact to affect scores. The biggest problems here are the use of judgmental and discretionary productivity points (and their impact), and dividing productivity points between organizational units.

(7) The system is not transparent. Most of the staff do not see a direct relationship between input effort and output score - which is one reason so many see the system as unfair.

(8) More time is required to determine the durability of the system. Part of the durability depends on lowering resistance to certain elements of the system (addressed in chapter 6).

b. In summary, several problems were found by the evaluation, but the system seems to be conceptually sound. It has more validity on the organizational level than for individuals, and probably has its greatest use over a period of several years so that short term fluctuations are averaged out.

5.2 CONCLUSIONS

5.2.1 General

a. TRASANA is a healthy organization with high morale, job satisfaction, and a motivated work force that is challenged and proud of the organization.

b. There is mixed reaction to TPMS. Supervisors have a more positive view than nonsupervisors, but the majority of the staff have legitimate concerns about TPMS fairness, and how TPMS will affect the creative, cooperative work environment.

c. TPMS is a conceptually sound productivity measurement system, but many of its functional elements need revision.

d. The project plan is a very positive contribution of TPMS.

e. Firm guidelines are needed on overhead so the system will accurately document the level of overhead use. Different categories of overhead are needed, especially in those situations where there is an overhead mission.

5.2.2 Productivity Index

a. The individual PI did not perform well and is the source of considerable staff uneasiness. Since accurate reflection of individual contributions in group analysis efforts cannot be captured with an objective measure, the system would be more accurate and less objectionable to the staff if the individual PI were eliminated.

b. The organizational PI would be greatly strengthened by eliminating overhead from the PI calculation.

c. The organizational PI will be a solid productivity indicator if the input estimates are more accurate. The PMB was not effective in controlling input estimates.

5.2.3 Product Quality Score

a. The PRB serves a necessary quality control function.

b. The quality scoring by the PRB is such a strong morale detriment that it should be changed.

c. PRB scoring procedures and authority should be more clearly defined and specified.

5.2.4 Scoring Algorithm

a. The scoring algorithm is unnecessarily complicated.

b. Judgmental and discretionary awards caused many problems and should be eliminated. They add a subjective element that is performance appraisal oriented, and did not work well. Individual performance appraisal and team productivity measurement should be separate.

c. Management flexibility is required to intervene in situations where circumstances beyond the team's control result in a poor score - such as undertaking a very risky project.

5.2.5 Follow-Up Evaluation

This study team has formed definite opinions about TPMS and has made substantial recommendations for changes in chapter 6. To preserve objectivity, any further TPMS evaluation should be performed by another study team. Since TPMS is a complex measurement system, a strong background in human measurement and management is needed.

CHAPTER 6

RECOMMENDATIONS

6.1 PURPOSE

This chapter contains recommendations for improving those aspects of TPMS which did not work well during the first scoring year.

6.2 APPROACH

The major problem areas are listed in table 6-1 along with the recommendations for correcting the problems. Specific strategies and procedures for implementing the recommendations are presented in appendix F in the form of an alternative TPMS scoring algorithm. The rationale behind each of the recommendations is discussed below.

6.3 RECOMMENDATIONS

6.3.1 Estimated Manpower (R_p)

a. The overestimation of manpower was the largest problem from the perspective of actually measuring productivity in the classical sense. Currently a division chief can enter a project into TPMS if his projected manpower requirement is not more than 125 percent of available manpower (125% rule). This 125 percent should be reduced to 110 percent to encourage more accurate estimating. Furthermore, it may need to be reduced even more in the future. The extra 10 percent is recommended for now because perfectly accurate estimating is not realistic.

b. The categorical scoring recommended for R, T, and Q will encourage more accurate manpower estimating by not rewarding overestimation. Furthermore, at the end of each scoring period each division's accuracy over the period can be determined, and a bonus given to the most accurate estimator. The bonus could take the form of increasing the division PI somewhat (the closer the factor score to 1.0, the higher the bonus), or some other reward could be devised. Since punishment is generally less effective than reward for effecting long term behavior change, and also has undesirable side-effects, rewarding the desired results is the favored approach over punishing inaccuracy.

6.3.2 Productivity Index (PI)

a. Individual. In its current form, the PI for individuals does not accurately reflect individual contributions and has considerable potential for abuse. An example would be a supervisor who accepts the PI as accurate without considering its limitations, and subsequently writes performance appraisals based on an "individual's productivity" using the PI. For this reason, retention of the individual PI is not recommended. A less desirable alternative is to eliminate judgmental and discretionary awards and assign every team member the team (project) score. The judgmental and discretionary awards

TABLE 6-1. RECOMMENDED SOLUTIONS TO TPMS PROBLEMS

PROBLEM	RECOMMENDATIONS
Estimated Manpower (R_p)	° Change 125% rule to 110% rule ° Institute categorical scoring which rewards accurate estimating ° Establish division level after-the-fact check with reward for accuracy (PI increase or bonus)
Productivity Index (PI)	
Individual	° Eliminate OR ° If retained eliminate judgmental and discretionary awards and each person gets the team PI for each project, weighted by manpower for aggregation
Organizational	° Eliminate overhead from PI calculation with a division level after-the-fact check and penalty for abuse (PI reduction) OR ° If overhead is retained in PI calculation, use different earning rates to avoid overhead mission penalty (e.g., mission overhead earns at the mean project rate for the year)
Quality Scoring (Q)	° Establish scoring sheet with scoring criteria to be used for each report ° Use three category scoring to eliminate fine score discriminations ° Assign productivity rating to PRB, which is determined in part by ratings from scored project study managers as to such factors as helpfulness, fairness, and thoroughness
Overhead	° Publish guidelines for charging overhead, including level to be considered acceptable without penalty, e.g. division staff and branch chiefs are primarily overhead ° Eliminate from PI calculations, except for abuse penalty (see organizational PI recommended solutions)
Algorithm	° Replace with scoring categories for R, T, and Q; eliminate C and P factors
Risk Taking	° Retain flexibility in system application - management interventions to avoid penalties from risky projects

are an attempt to make an organizational productivity measurement system serve as an individual performance appraisal device. As shown in the first year results, this just did not work. The system will work better without the performance appraisal aspect (be a stronger organizational index), and there is already a separate personnel appraisal system where supervisors can reflect individual job performance.

b. Organizational. On the organizational level, the problem is overhead. One approach is to eliminate overhead from the PI computation, and to have a penalty for abuse of overhead. Another approach is to have different earning rates for overhead (see 6.3.4).

6.3.3 Quality Scoring (Q)

The PRB serves a very valuable quality control function which should be retained; however, there are several difficulties with the current Q scoring. Categorical scoring (excellent, okay, revision required) will help considerably, but that is not enough. A checks-and-balance procedure is needed to avoid the possibility of the PRB exerting too much influence over reports. The primary responsibility for reports should rest with line management. The PRB should be required to complete a checklist and rating scale for each report which is promptly furnished to the responsible division chief. This checklist should contain those factors the board uses in determining the quality score. It is likely that these factors are to a large extent report organization and readability.[1] Additionally, accountability is a two way street. Permanent PRB members (voting members) should be held accountable for their productivity by being scored on their thoroughness, helpfulness, and fairness by project study managers, branch chiefs, and division chiefs (instances were reported of PRB members not having read the report prior to the board meeting). A simple one-page rating form for this purpose could be developed rather easily.

6.3.4 Overhead

Guidelines for charging overhead are essential if this problem is to be overcome. The guidelines should establish "acceptable" overhead limits. For example, supervisors are generally overhead, and there are legitimate overhead activities that take up a small amount of everyone's time. This is widely recognized in such activities as applying for research grants, where a figure around 30 percent is fairly typical of expected research overhead expense. The first year treatment of overhead diluted PI usefulness.

---

[1] *The Colonel who served as PRB board president during the first scoring period was interviewed before departing TRASANA. He reluctantly admitted that this was the case even though he "wished" it could different. He had no recommendations for how to change the situation. The problem is that PRB members do not have enough time to gather and process detailed information on each and every project. The possible return on the investment of having them take such time would not likely be worth it.*

6.3.5 Scoring Algorithm

If three categories are used for the R, T, and Q factors, the current algorithm can be replaced by 27 possible outcomes without loss of information, and with reduced record keeping requirements (especially if individual PI is discontinued). These 27 possible outcomes should be evaluated and weighted in advance, so a project score is automatically assigned based on the outcome category (see appendix F for details).

6.3.6 Risk Taking

There is considerable concern that TPMS will reduce staff willingness to take on unknown and risky projects. Since no workable risk index is available, this problem can be handled by not applying TPMS scoring dogmatically. Management should reserve the perogative of adjusting project scores where risk resulted in inequitable results.

APPENDIX A

TPMS-TE STUDY PLAN

STUDY PLAN FOR THE PHASE I

EVALUATION OF TRASANA TECHNICAL

STAFF PRODUCTIVITY MEASUREMENT

SYSTEM-TEST (TPMS-TE)

MAY, 1981

TABLE OF CONTENTS

STUDY PLAN FOR THE PHASE I EVALUATION OF TRASANA TECHNICAL STAFF PRODUCTIVITY MEASUREMENT SYSTEM-TEST (TPMS-T)

1. PURPOSE

To evaluate Phase I (the first year) of the three-year test implementation of the TRASANA Technical Staff Productivity Measurement System (TPMS).

2. BACKGROUND

a. TPMS is a methodology designed to measure the productivity of the TRASANA technical staff, its organizational elements, and its individual employees. It yields a single dimensionless index as the measure of productivity. The formulation of this index is intended to be consistent with the traditional definition of productivity as value of output divided by value of input. In principle, TPMS places TRASANA on a "fixed price plus incentive basis for producing its various products. TPMS was designed with the following attributes in mind:

(1) Relevant - to the organization, its products, and to the jobs that produce the products.

(2) Rational and fair - so that it can be accepted and used by a sophisticated workforce.

(3) Affordable - in terms of direct and indirect/overhead costs for operation and maintenance.

(4) Flexible - to couple with a dynamic workload, organization, and operations.

(5) Simple - both in principle and practice.

(6) Transparent - so as to facilitate understanding of cause and effect relationships.

(7) Durable - to support multi-year use.

b. Given that TPMS enables reasonably equal opportunity for productivity and scoring throughout the organization, then the resulting productivity index (PI) can be used to compare the relative productivity of TRASANA divisions, branches, and individuals. Management and individual goals can be structured in terms of changes in the PI. The degree to which the goals are met could be a factor in performance appraisal. The TPMS has been adopted at TRASANA for a three year trial beginning 1 Oct 80. This study is the first phase of a 3 year, 3 phase, evaluation of the TPMS test (TPMS-T).

3. PROBLEM

To determine the degree to which TPMS provides a meaningful measure of the productivity of the TRASANA technical staff and its organizational elements.

4. IMPACT OF PROBLEM

The purpose of the TPMS test period is to develop experience working with the system so its effects on management practices, employee motivation, and productivity can be judged. A systematic collection and assimilation of information about the system test is required to assess these factors. Implementation of an untested system could result in having inaccurate productivity results, or even hamper productivity by adversely affecting management practices and employee motivation.

5. SCOPE

a. The study will be limited to evaluating the results of the first trial implementation year of TPMS.

b. TPMS was developed in part as a local response to requirements of the new Performance Appraisal (PA) and Merit Pay (MP) systems. While informal integration of TPMS data and PA and MP systems may occur during the test period, this evaluation is limited exclusively to TPMS. No assessment of the PA or MP systems is intended.

c. TPMS-TE is an assessment of a specific productivity measurement system. It is not intended to be a comprehensive study of methods to measure TRASANA or research and development type productivity.

d. One of the desired characteristics of TPMS is that it be affordable. Available data will be collected on resources expended on the first year test of TPMS; however, judgments on whether the costs are a worthwhile investment are beyond the scope of this study.

6. OBJECTIVES FOR PHASE I EVALUATION

a. To determine how the TRASANA technical staff (TTS) perceives TPMS. The study questions for this objective are:

(1) What is the TTS general attitude toward TPMS?

(2) What is the perceived motivational impact of TPMS?

(3) Is TPMS seen as a viable managerial tool?

b. To determine the degree to which TPMS is a meaningful measure of productivity. The study questions for this objective are:

(1) To what extent do TPMS ratings agree with confidential supervisory productivity assessments for individuals and branches.

(2) Does TPMS have generally the same impact on the different TRASANA organization divisions?

(3) What are the functional relationships among the scoring algorithm variables?

(4) What is the operational impact of the Productivity Measurement Board (PMB) and Product Review Board (PRB)?

c. To develop a TPMS data base consisting of Phase I TPMS-T results. Available data on resource implications will be included. This data base will form the benchmark for comparison of future results.

7. ASSUMPTIONS

a. Phase I of the TPMS-T implementation will end 30 June 1981.

b. TPMS will remain essentially unchanged during Phase II of the test implementation (test integrity).

c. The test will apply equally to all TRASANA Technical Staff (e.g. civilian and military).

d. Key organizational supervisors know the relative long term productivity of employees they supervise.

e. Division chiefs, branch chiefs, and the TTS will provide complete, accurate and timely information for the study.

f. First year test results and other required inputs will be provided on time, including manpower and computer time charges.

g. Management will respect the confidential data collected as a part of the study.

8. CONSTRAINTS

a. Not all of the TTS will have completed TPMS scored projects during the first test implementation year.

b. There is no basis for comparing test period productivity with past productivity.

9. DATA COLLECTION METHODOLOGY

The study methodology consists of several independent data collection steps and the development of accompanying data collection instruments.

a. TRASANA Technical Staff Perception Questionnaire.

(1) This questionnaire will be designed to gather the data required to answer the three study questions supporting the first objective: general attitude toward TPMS, perceived motivational aspects of TPMS, and the perceived impact of using TPMS as a managerial tool. After the questionnaire is constructed it will be tested, evaluated for reliability, and revised as appropriate before administering to the TRASANA staff.

(2) The final version questionnaire will be administered to the available TTS at the end of the first test year, but before final first year PIs are compiled (July 1981). This timing is to allow maximum exposure to the system, but to gather perceptions before the first year results have had a chance to impact impressions.[1] The surveys will be administered on a group basis to minimize interaction of staff members while responding.

b. Productivity Criterion Measure. Based on a literature survey of research and development productivity factors, a short scale or index will be devised for supervisory personnel (division and branch chiefs) to use in evaluating technical employee productivity. This scale or index will be used in interviews to obtain a supervisory assessment of productivity. This assessment will be conducted privately, and without the individual employee or any other person being given access to the ratings except the analysts assigned to this project.[2] Division chiefs will rate branch chiefs, and branch chiefs will rate persons in their branches. Because of the sensitivity of this information, special precautions will be employed. Ratings obtained for this study will not be available to supervisory staff, will not be placed in personnel files, and will not be used in any way that would allow individual identification.

c. Survey of Divisional Impact. Survey questions on relevant issues, such as changes in managerial practices, will be developed and distributed in advance to division chiefs. Available division chiefs will be interviewed to obtain answers to the impact questions and any other information relevant to the TPMS test evaluation.

d. TPMS-T first year results. Project scoring sheets, productivity points (PP), and productivity indices will be collected for each individual, branch, and division.

e. PMB and PRB functioning. Rating forms will be developed for surveying PMB members on its functioning, and to survey study managers on the PRB functioning. All available PMB members will be surveyed, and study managers who have completed scored projects under TPMS-T will be sampled.

10. DATA ANALYSIS METHODOLOGY

The data analysis methodology employed to answer the study questions and accomplish the objectives are now described. Table 1 presents the objectives and study questions by data sources and principal analytic techniques.

---

[1]*At this time it is unknown how the first year results will be distributed and what emphasis will be put on them in the different divisions and branches. To avoid the possible contaminating effects of first year scores, the perception survey will be administered before the final scores are available.*

[2]*Currently only two analysts are involved in this project. A third analyst will be assigned eventually to work on data analysis. It is not anticipated that the third analyst will have access to this information until it is coded in such a manner as to preclude individual identification.*

TABLE 1

DATA SOURCES AND ANALYTIC TECHNIQUES FOR

ANSWERING STUDY QUESTIONS AND OBJECTIVES

OBJECTIVE/STUDY QUESTION	DATA SOURCES	ANALYTIC TECHNIQUES
I. TTS perceptions of TPMS		
(1) General Attitude	Q, I	Descriptive statistics; inferential statistics to determine differences among organizational units.
(2) Perceived Motivational Aspects	Q, I	
(3) Perception as Management Tool	Q, I	
II. Meaningful measure of productivity		
(1) Relationship between PI and supervisory ratings	C, T	Inferential statistics including correlation analysis.
(2) Equivalent divisional impact	Q, I, T	Descriptive statistics; inferential statistics to determine differences among divisions.
(3) Relationships among PI Variables	T	Mathematical description of the scoring algorithm.
(4) Impact of PMB and PRB	Q, I, T, S, B	Descriptive statistics.

Q - Questionnaire

I - Division chief interviews

T - TPMS first implementation year data (PP, PI, project score sheets, project plans)

C - Criterion measure (division and branch chiefs)

S - Study manager survey

B - PMB member survey

a. TTS perceptions of TPMS. This includes study questions on general attitude, motivational aspects, and perceptions as a management tool. Data from the TTS questionnaire and division chief interviews will be analyzed using descriptive statistics. Inferential statistics will be used to test comparisons by organizational units and demographic characteristics.

b. Relations between rankings by PI and rankings by supervisors. Inferential statistics including correlation analysis will be used to determine the relationship between ranks of individuals by the PI and supervisors.

c. Impact on divisions. Data from the TTS questionnaire, division chief interviews, and first year scoring data will be analyzed using descriptive statistics. Comparison of TPMS scores by divisions will be made. Inferential statistics will be employed to test the significance of any differences among divisions.

d. Relationships among the scoring algorithm variables. The scoring algorithm will be analyzed mathematically. The relationships among the variables will be described and their relative effect on the PI assessed.

e. Impact of PMB and PRB. Relevant data from the TTS questionnaire, division chief interviews, first year scoring results, study manager survey, and PMB survey will be analyzed descriptively. PMB effectiveness in reviewing project plans and critiquing resource estimates will be assessed. Likewise the operational impact of the PRB will be assessed.

f. Table 2 summarizes the project tasks including data collection and data analysis procedures. For a more detailed description of individual tasks, see the Phase I Evaluation of TRASANA Technical Staff Productivity Measurement System (TPMS)-Test Project Plan, 29 April 1981.

11. OPERATIONAL SUPPORT REQUIREMENTS

a. Deputy Director for Technical Operations

(1) Approve project plan and study plan.

(2) Guarantee confidentiality of sensitive data.

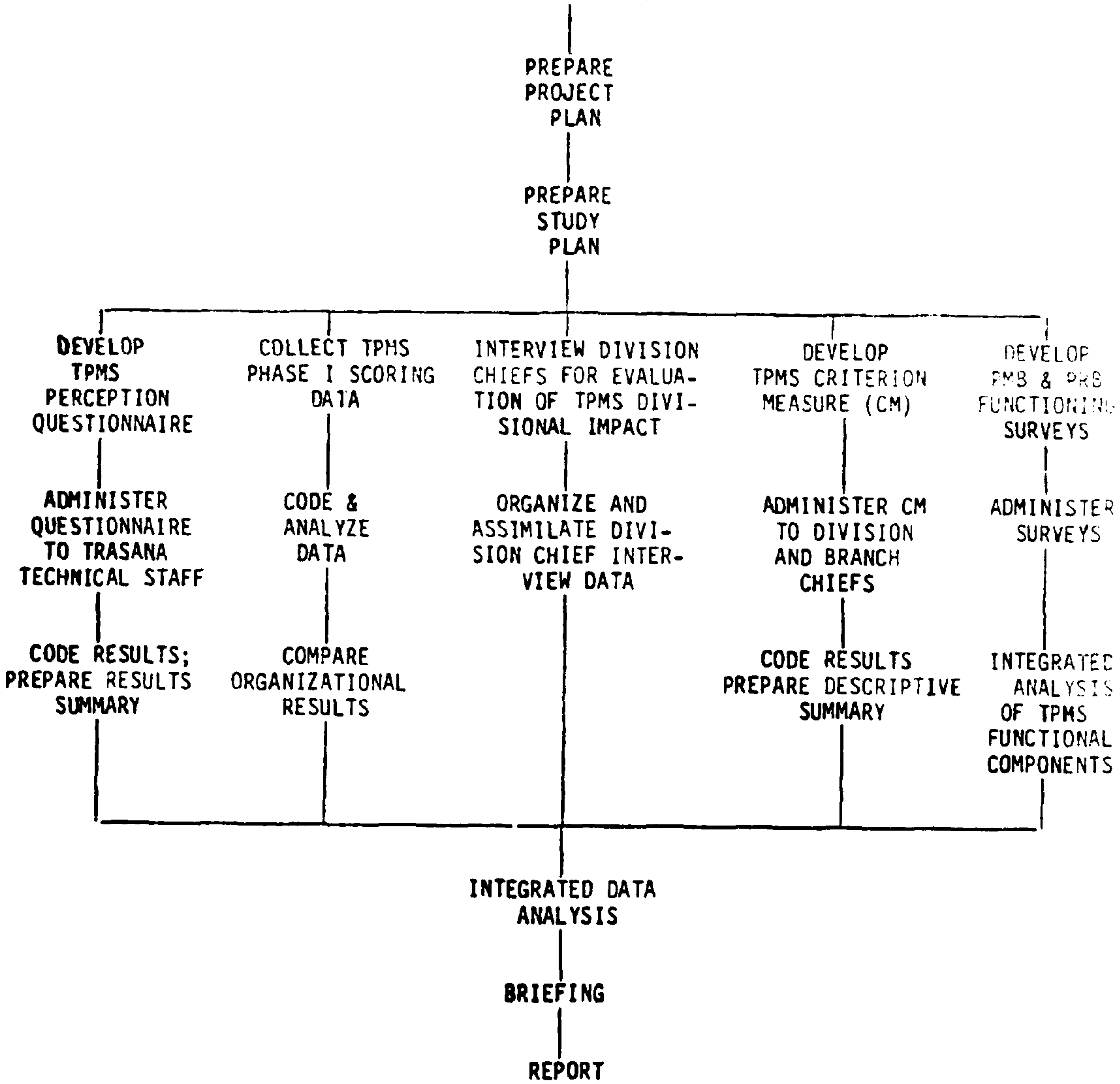
(3) Maintain TPMS-T test integrity.

(4) Provide accessibility of study team to division chiefs, branch chiefs, TRASANA technical staff, PMB and PRB participants and procedures.

(5) Provide support in the coordination and execution of data collection activities (including assembly of necessary staff for surveys).

(6) Encourage full cooperation of all TRASANA supervisors and technical staff with the TPMS-T evaluation effort.

TABLE 2
TPMS-TE TASKS



(7) Provide needed support and assistance to help study team overcome unforeseen obstacles.

b. Division Chiefs

(1) Participate in data collection including individual productivity assessments of branches.

(2) Provide an assessment of the impact of TPMS on division managerial practices, productivity, and employee morale.

(3) Afford study team free and unrestrained access to division personnel for the required data collection procedures.

(4) Encourage full cooperation of employees with study effort.

c. Branch Chiefs

(1) Participate in data collection including individual productivity assessment of branch members.

(2) Afford study team free and unrestricted access to branch personnel for the required data collection procedures.

(3) Encourage full cooperation of employees with study effort.

d. Study Team

(1) Develop project plan and study plan.

(2) Develop study instruments.

(3) Collect required data.

(4) Conduct study in accordance with study plan and project plan.

(5) Safeguard sensitive information collected as part of this study.

(6) Keep Deputy Director of Technical Operations informed on study progress, including a briefing of Phase I results by 23 December 1981.

(7) Prepare a report of Phase I results.

12. SCHEDULE

Milestone	Date
a. Draft Project Plan	30 Apr 81
b. Draft Study Plan	1 Jun [illegible]
c. TRASANA Technical Staff Survey	[illegible]
d. TRMS First Year Results Collected	[illegible] 81
e. Criterion Measure Data Collected	4 Sep 81
f. Division Chief Interviews Completed	16 Oct 81
g. Data Analysis Completed	1 Dec 81
h. Brief Results to Deputy Director for Technical Operations	23 Dec 81
i. Phase I Report Completed	29 Jan 82

APPENDIX B

TRASANA TECHNICAL STAFF SURVEY

TRASANA TECHNICAL STAFF SURVEY

TURN THE PAGE AND READ THE INFORMATION

AND INSTRUCTIONS

PURPOSE

The purpose of this questionnaire is to find out how the TRASANA Technical Staff views the TRASANA Productivity Measurement System (TPMS) at this time.

GENERAL INFORMATION

This questionnaire has been designed to answer specific management questions about TPMS as a part of the evaluation of the TPMS three year trial. Honest answers to the questions are required for a meaningful evaluation. You are not required to give your name, but your organizational unit (office symbol) is needed for grouping responses. Results will be presented to management in the form of summary statistics only.

This survey has three parts. Part I asks general background information. Part II consists of several attitude statements about TRASANA and TPMS. You are to indicate on the answer sheet the extent to which you agree or disagree with each statement. You will probably disagree with some statements and agree with others. We are interested in the extent to which you agree or disagree with the stated opinions. Part III has three questions on your familiarity with TPMS. The entire survey should take approximately 15 minutes.

GENERAL INSTRUCTIONS

1. Use a soft lead pencil for marking the answer sheet.

2. Read the instructions for each part of the questionnaire.

3. Read each item carefully, and be sure the number on the answer sheet matches the item you are answering.

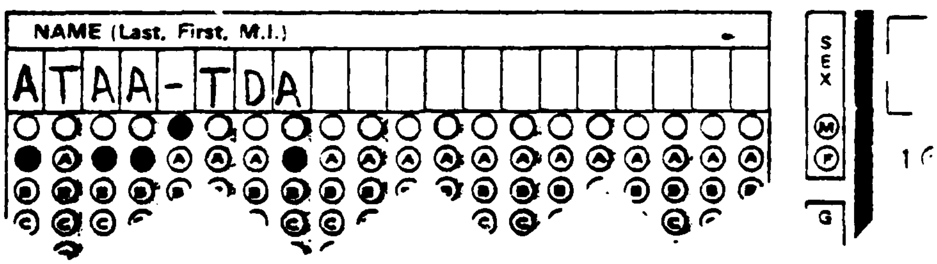
4. Fill in the circle on the answer sheet which BEST describes your response to the statement. Select only one response for each statement.

5. If you make a mistake or wish to change your answer, erase the mark completely before entering a new one.

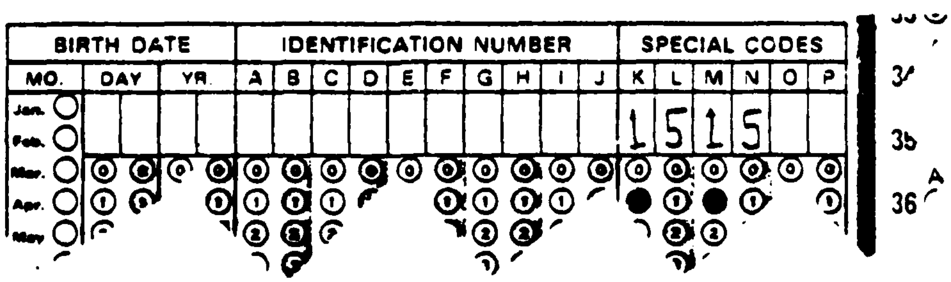
6. Answer every item. If the answers provided are not exactly what you would like, select the one which is closest.

ANSWER SHEET

1. In the space "NAME", print the complete office symbol for your branch. If you are assigned to a divisional level job, enter your complete office symbol. EXAMPLE:



2. Under "SPECIAL CODES" enter (left justified) the four number designation for your career specialty area if you are a civilian. Military personnel enter duty MOS. EXAMPLE:



3. The last page of this booklet is for you to add any comments about TPMS or the three year test of TPMS (TRASANA MEMO 5-4) which you feel would be helpful to the test evaluation, including suggestions for improvements.

PART I: BACKGROUND INFORMATION

INSTRUCTIONS

Answer all questions. Mark your answers on the answer sheet provided with a soft lead pencil. Select only one answer for each question.

1. I am a (an)

   A. Civilian
   B. Active duty officer
   C. Active duty enlisted soldier

2. My age is

   A. Under 30
   B. 30 thru 39
   C. 40 thru 49
   D. 50 thru 59
   E. 60 thru 64
   F. 65 thru 69
   G. 70 and over

3. My grade is (If not civilian mark "I")

   A. GS-07
   B. GS-09
   C. GS-11
   D. GS-12
   E. GS-13
   F. GS-14
   G. GS-15
   H. Other
   I. Not a civilian

4. My military pay grade is (If not active duty officer mark "G")

   A. 0-1
   B. 0-2
   C. 0-3
   D. 0-4
   E. 0-5
   F. 0-6
   G. Not active duty officer

5. My military pay grade is (If not active duty enlisted soldier mark "G")

A. E-4
B. E-5
C. E-6
D. E-7
E. E-8
F. E-9
G. Not active duty enlisted soldier

6. I have been in my present grade for

A. Less than 1 year
B. 1 year but not 2
C. 2 years but not 3
D. 3 years but not 4
E. 4 years but not 5
F. 5 years but not 6
G. 6 years but not 7
H. 7 years but not 8
I. 8 years but not 9
J. 9 years or more

7. I have worked for TRASANA (or its SAFEGUARD predecessor) for

A. Less than 1 year
B. 1 year but not 2
C. 2 years but not 3
D. 3 years but not 4
E. 4 years but not 5
F. 5 years but not 6
G. 6 years but not 7
H. 7 years but not 8 (SAFEGUARD)
I. 8 years but not 9 (SAFEGUARD)
J. 9 years or more (SAFEGUARD)

8. I have worked for the Federal Government (including military service) for a total of

A. Less than 1 year
B. 1 thru 3 years
C. 4 thru 9 years
D. 10 thru 19 years
E. 20 thru 29 years
F. 30 years or more

9. The highest academic degree or diploma I have is

A. High School Diploma/GED
B. Associate degree
C. Bachelor's degree
D. Master's degree
E. Doctor's degree

10. My primary academic specialty, or the area in which I have had the highest level of advanced study, is

A. Engineering
B. Physical sciences (physics, chemistry, etc.)
C. Mathematics/statistics
D. Computer science
E. Social sciences (psychology, sociology, etc.)
F. Education
G. Business (management, accounting, etc.)
H. Liberal arts (music, art, history, etc.)
I. Other

11. My position is best described as

A. Clerical
B. Technical/Professional (team chief, team member, support)
C. Advisory (non-supervisory)
D. First-line supervisor (Branch Chief)
E. Second-line supervisor (Division Chief or Deputy)
F. None of the above

12. I am a

A. Male
B. Female

13. My primary ethnic identity is

A. American Indian
B. Hispanic
C. Afro-American
D. Asian American
E. Other

GO ON TO PART II

PART II: TPMS OPINION SURVEY

INSTRUCTIONS

This section of the survey contains a series of attitude statements designed to determine how the staff feels about TPMS and its relation to the type of work we do at TRASANA. Read each statement, and decide whether you agree or disagree with it. Next, select the letter that BEST describes to what extent you agree or disagree with the statement, and mark the circle under the letter on the answer sheet. In the event that none of the six responses adequately reflects your opinion on the statement, select the one which is closest to the way you feel, or which is least objectionable to you. Answer all questions.

CHECK TO BE SURE YOU ARE AT THE RIGHT PLACE ON THE ANSWER SHEET.

	Agree: Strongly	Agree: Somewhat	Agree: Slightly	Disagree: Slightly	Disagree: Somewhat	Disagree: Strongly
14. Overall, the performance standards at TRASANA are high.	A	B	C	D	E	F
15. I am proud to be associated with TRASANA because of its quality products.	A	B	C	D	E	F
16. I like working at TRASANA.	A	B	C	D	E	F
17. There is a lot of "dead wood" at TRASANA.	A	B	C	D	E	F
18. Most of the people at TRASANA do high quality work.	A	B	C	D	E	F
19. The management at TRASANA does its best to provide good working conditions.	A	B	C	D	E	F
20. Supervisors in TRASANA will take care of their people, with or without TPMS.	A	B	C	D	E	F
BE SURE YOU ARE AT 21 ON THE ANSWER SHEET						
21. Most people are treated fairly at TRASANA.	A	B	C	D	E	F
22. The overall productivity of TRASANA is not really very high.	A	B	C	D	E	F
2 . If TRASANA were on a profit basis, such as a private contractor or consulting firm, it would <u>not</u> be making money.	A	B	C	D	E	F
24. Productivity in my branch often suffers from inadequate organization and planning.	A	B	C	D	E	F
25. Productivity in my division often suffers from inadequate organization and planning.	A	B	C	D	E	F
26. I like doing the kind of work I do at TRASANA.	A	B	C	D	E	F

	Agree			Disagree		
	Strongly	Somewhat	Slightly	Slightly	Somewhat	Strongly
27. My job is not very challenging.	A	B	C	D	E	F
28. I am generally satisfied with my job performance.	A	B	C	D	E	F
29. I do not get much personal satisfaction from my job.	A	B	C	D	E	F
30. The skills I have are well matched to the requirements of my job.	A	B	C	D	E	F
BE SURE YOU ARE AT 31 ON THE ANSWER SHEET						
31. I do not get much useful feedback on how I am doing in my job.	A	B	C	D	E	F
32. My supervisors generally give me the respect I deserve.	A	B	C	D	E	F
33. My supervisors do not know what I can do.	A	B	C	D	E	F
34. My supervisor does a good job of planning work and managing people.	A	B	C	D	E	F
35. I am usually free to decide on my own how to go about accomplishing the tasks I am given to do.	A	B	C	D	E	F
36. I usually have very little say in the schedule for completing my tasks.	A	B	C	D	E	F
37. I can use a lot of initiative and try new things in my job.	A	B	C	D	E	F
38. My job typically involves doing a small part of a large project for which someone else is reponsible.	A	B	C	D	E	F
39. Quantifying productivity in the kind of work I do can be done with reasonable accuracy (i.e. distinguishing between the quality of contribution made by different people over a period of time).	A	B	C	D	E	F

	Agree: Strongly	Agree: Somewhat	Agree: Slightly	Disagree: Slightly	Disagree: Somewhat	Disagree: Strongly
40. TRASANA productivity can be increased substantially without hiring more people.	A	B	C	D	E	F
41. Most people are giving TPMS a fair chance.	A	B	C	D	E	F
42. TPMS is a simple, easy to understand system.	A	B	C	D	E	F
43. TPMS is flexible enough to accommodate most of the projects done in TRASANA.	A	B	C	D	E	F
44. TPMS places emphasis on the factors that reflect differences in productivity.	A	B	C	D	E	F
45. The distribution of productivity points will show how the study manager views the work quality of individual team members.	A	B	C	D	E	F
46. Productivity points and indices do not really matter right now because the system is just being tested.	A	B	C	D	E	F
47. With experience, manpower resources required for projects can be realistically estimated.	A	B	C	D	E	F
48. With experience, computer resources required for projects can be realistically estimated.	A	B	C	D	E	F
49. I have seen some positive effects of TPMS so far.	A	B	C	D	E	F
50. In my division, most people have made up their minds about TPMS.	A	B	C	D	E	F
51. I do not expect to score very well under TPMS.	A	B	C	D	E	F
52. Considering that measuring productivity is not easy, TPMS is basically a fair system.	A	B	C	D	E	F
53. In the long run, all divisions will have an equal chance to earn productivity points under TPMS.	A	B	C	D	E	F

	Agree			Disagree		
	Strongly	Somewhat	Slightly	Slightly	Somewhat	Strongly
54. TPMS is objective.	A	B	C	D	E	F
55. People who produce the best work will generally get the most productivity points under TPMS.	A	B	C	D	E	F
56. TPMS can be easily manipulated.	A	B	C	D	E	F
57. I understand the mechanics of TPMS.	A	B	C	D	E	F
58. Most of the technical staff do not understand TPMS.	A	B	C	D	E	F
59. I know the five factors used in calculating a project score.	A	B	C	D	E	F
60. I do not know how a project score is computed.	A	B	C	D	E	F
TURN ANSWER SHEET OVER						
61. I understand the difference between productivity points (PP) and productivity index (PI).	A	B	C	D	E	F
62. Given all the input factor values, I could compute my own productivity index right now.	A	B	C	D	E	F
63. As long as a person does his job, he need not worry about TPMS because he will come out okay.	A	B	C	D	E	F
64. For the most part, productivity points are distributed according to the amount of time spent on a project.	A	B	C	D	E	F
65. If a division chief feels an individual has not been justly rewarded for work on a project, he can create extra productivity points for that person.	A	B	C	D	E	F

	Agree: Strongly	Agree: Somewhat	Agree: Slightly	Disagree: Slightly	Disagree: Somewhat	Disagree: Strongly
66. A person has the option of not using time charged to overhead activities when calculating his productivity index.	A	B	C	D	E	F
67. Under TPMS, a study manager is encouraged to work for the highest quality (PRB) score, even at the expense of being late or over on manpower.	A	B	C	D	E	F
68. Productivity points and merit pay are not related at this time.	A	B	C	D	E	F
69. TPMS will only affect management.	A	B	C	D	E	F
70. Management practices have been changed by TPMS.	A	B	C	D	E	F
BE SURE YOU ARE AT 71 ON THE ANSWER SHEET						
71. More planning is being done for projects now than before TPMS.	A	B	C	D	E	F
72. TPMS will be a useful tool for management.	A	B	C	D	E	F
73. After the bugs are worked out, TPMS will be a useful factor for individual performance appraisals.	A	B	C	D	E	F
74. My productivity index (PI) under TPMS is pretty much beyond my control.	A	B	C	D	E	F
75. My supervisor has discussed with me the distribution of productivity points on the projects on which I have worked.	A	B	C	D	E	F
76. Management practices have improved as a result of TPMS.	A	B	C	D	E	F
77. TPMS represents no major change in emphasis by TRASANA management.	A	B	C	D	E	F

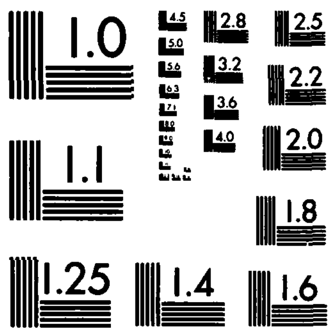
AD-A122 285 PHASE I EVALUATION OF TRASANA TECHNICAL STAFF PRODUCTIVITY MEASUREMENT SY..(U) ARMY TRADOC SYSTEMS ANALYSIS ACTIVITY WHITE SANDS MISSILE RAN.. JAN 82 2/2

UNCLASSIFIED TRASANA-TR-5-82 SBI-AD-F050 072 F/G 5/9 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

	Agree			Disagree		
	Strongly	Somewhat	Slightly	Slightly	Somewhat	Strongly
78. TPMS will motivate people to work harder.	A	B	C	D	E	F
79. My performance appraisal for this year will not be influenced by TPMS.	A	B	C	D	E	F
80. TPMS will not affect me very much.	A	B	C	D	E	F
81. Assuming that productivity can be increased by doing more work or doing higher quality work, TPMS will lead to increased productivity in the long run.	A	B	C	D	E	F
82. TPMS has resulted in a more competitive environment among the TRASANA staff.	A	B	C	D	E	F
83. If TPMS is adopted at TRASANA, things will not really change very much.	A	B	C	D	E	F
84. TPMS has substantially affected my branch.	A	B	C	D	E	F
85. If TPMS is adopted at TRASANA, productivity will not change.	A	B	C	D	E	F
86. I do not care whether TPMS succeeds or fails in its three year test.	A	B	C	D	E	F
87. I do not know very much about TPMS.	A	B	C	D	E	F
88. I would like to know more about TPMS.	A	B	C	D	E	F

GO ON TO PART III

PART III: TPMS FAMILIARITY AND COMMENTS

INSTRUCTIONS

This section of the survey contains questions designed to find out how familiar the TRASANA staff is with various aspects of TPMS at this time. Read each question carefully and record your answers on the answer sheet. Answer all questions

CHECK YOUR ANSWER SHEET TO BE SURE YOU ARE ON THE CORRECT NUMBER

89. Which of the following best describes your familiarity with TRASANA Memorandum 5-4?

A. I do not know what it is.
B. I know what it is but have not seen it.
C. I have seen it, but have not looked through it.
D. I have browsed through it quickly.
E. I have read it.
F. I have read it more than once.
G. I have studied it thoroughly.

90. Which of the following best describes your overall understanding of TPMS at this time?

A. I fully understand it.
B. I understand it fairly well.
C. I do not understand some parts of it.
D. I understand very little of it.
E. I do not know anything about it.

91. Which of the following statements best describes your overall feeling about TPMS at this time?

A. I like it.
B. There are some good and bad points to it, but the good probably outweigh the bad.
C. There are some good and bad points to it, but the bad probably outweigh the good.
D. I don't like it.

Please use the attached page for comments about TPMS which would be helpful to the evaluation team, including suggestions for changing TPMS or 5-4. Thank you.

APPENDIX C

SUPPORTING TABLES

TABLE C-1

TTS AND SURVEY SAMPLE BY DIVISION

DIVISION	PERCENT OF TTS	PERCENT OF SAMPLE
A	12.6	11.1
B	9.5	10.5
C	13.0	12.6
D	15.0	15.8
E	14.6	17.9
F	11.1	7.4
G	8.3	6.3
H	15.8	18.4
TOTAL	99.9[a]	100

[a] *Does not total 100% due to rounding.*

TABLE C-2
AGE OF RESPONDENTS

AGE CATEGORY	NUMBER	PERCENT
Under 30	21	11.1
30 to 39	57	30.0
40 to 49	78	41.1
50 to 59	29	15.3
60 to 64	4	2.1
65 to 69	1	0.5
TOTAL	190	100.1[a]

[a]*Does not total 100% due to rounding.*

TABLE C-3
RESPONDENT TIME AT TRASANA

TIME (YEARS)	NUMBER	PERCENT
Less than 1 yr	25	13.2
1	21	11.1
2	8	4.2
3	3	1.6
4	7	3.7
5	11	5.8
6	19	10.0
7	0	0
8	6	3.2
9 or more	90	47.4
TOTAL	190	100.1[a]

[a]*Does not total 100% due to rounding.*

TABLE C-4
RESPONDENT TIME IN GRADE

TIME (YEARS)	NUMBER	PERCENT
Less than 1 yr	31	16.3
1	17	8.9
2	12	6.3
3	16	8.4
4	16	8.4
5	20	10.5
6	10	5.3
7	8	4.2
8	6	3.2
9 or more	54	28.4
TOTAL	190	99.9[a]

[a]*Does not total 100 due to rounding.*

TABLE C-5
RESPONDENT TIME IN FEDERAL SERVICE

TIME (YEARS)	NUMBER	PERCENT
Less than 1 yr	3	1.6
1 to 3	13	6.8
4 to 9	36	18.9
10 to 19	75	39.5
20 to 29	60	31.6
30 or more	3	1.6
TOTAL	190	100

TABLE C-6

HIGHEST ACADEMIC DEGREE OF RESPONDENTS

DEGREE	NUMBER	PERCENT
High School/GED	8	4.2
Associate	4	2.1
Bachelor	91	47.9
Master	79	41.6
Doctor	8	4.2
TOTAL	190	100

TABLE C-7

PRIMARY ACADEMIC SPECIALTY OF RESPONDENTS

SPECIALTY	NUMBER	PERCENT
Mathematics/Statistics	56	29.5
Engineering	54	28.4
Physical Sciences	32	16.8
Computer Science	12	6.3
Social Sciences	11	5.8
Education	9	4.7
Business	7	3.7
Liberal Arts	3	1.6
Other	6	3.2
TOTAL	190	100

TABLE C-8

TECHNICAL STAFF SURVEY RESPONSES TO
ITEMS 14 TO 88
(PERCENTAGES)

ITEM NUMBER	AGREE STRONGLY A	AGREE SOMEWHAT B	AGREE SLIGHTLY C	DISAGREE SLIGHTLY D	DISAGREE SOMEWHAT E	DISAGREE STRONGLY F
14	27	48	16	6	2	1
15	33	38	19	6	3	1
16	53	29	12	2	3	1
17	16	22	24	16	13	9
18	22	45	20	9	3	1
19	26	31	22	6	8	7
20	26	35	20	9	5	5
21	28	38	21	6	6	1
22	5	14	21	17	25	18
23	18	12	18	14	23	15
24	9	14	21	12	19	25
25	5	20	24	16	16	19
26	49	27	13	5	3	3
27	6	8	11	8	27	40
28	33	40	16	6	3	2
29	5	15	12	14	23	31
30	30	39	11	8	7	5

TABLE C-8 (CONTINUED)

ITEM NUMBER	AGREE STRONGLY A	AGREE SOMEWHAT B	AGREE SLIGHTLY C	DISAGREE SLIGHTLY D	DISAGREE SOMEWHAT E	DISAGREE STRONGLY F
31	15	21	18	12	19	15
32	39	33	14	5	3	6
33	7	12	17	11	30	23
34	23	36	19	7	9	6
35	51	27	14	4	2	2
36	12	11	12	18	24	23
37	39	31	14	7	5	4
38	19	19	12	8	17	25
39	5	14	17	11	22	31
40	22	29	25	6	13	5
41	10	33	38	7	8	4
42	3	15	15	18	22	27
43	4	21	18	20	18	20
44	1	13	26	20	14	26
45	7	21	23	13	17	19
46	4	11	16	21	19	29
47	10	39	20	12	9	10
48	8	30	23	16	12	11
49	4	12	18	18	18	30

TABLE C-8 (CONTINUED)

ITEM NUMBER	AGREE STRONGLY A	AGREE SOMEWHAT B	AGREE SLIGHTLY C	DISAGREE SLIGHTLY D	DISAGREE SOMEWHAT E	DISAGREE STRONGLY F
50	12	24	24	29	7	5
51	12	12	19	19	24	14
52	3	15	32	15	16	19
53	4	11	22	17	23	23
54	4	9	28	17	13	28
55	4	14	21	18	17	26
56	27	33	23	11	6	0
57	17	37	21	14	6	5
58	5	16	22	26	22	9
59	31	21	18	7	7	16
60	12	7	14	18	17	32
61	28	23	16	10	7	16
62	23	20	12	14	8	23
63	2	19	12	23	21	23
64	14	35	20	15	10	6
65	17	14	21	17	13	18
66	14	8	19	18	19	22
67	8	12	12	19	23	26
68	28	23	17	12	10	10

TABLE C-8 (CONTINUED)

ITEM NUMBER	AGREE STRONGLY A	AGREE SOMEWHAT B	AGREE SLIGHTLY C	DISAGREE SLIGHTLY D	DISAGREE SOMEWHAT E	DISAGREE STRONGLY F
69	2	3	5	12	22	56
70	14	22	28	13	11	12
71	12	29	31	9	8	11
72	9	25	30	10	12	14
73	6	17	24	12	22	19
74	16	20	21	19	15	9
75	18	13	13	7	11	38
76	1	8	24	24	18	25
77	9	14	19	24	22	12
78	1	12	26	15	19	27
79	16	14	18	28	16	8
80	7	8	13	20	26	26
81	12	17	22	14	16	19
82	11	20	23	20	13	13
83	5	14	24	20	18	19
84	5	14	29	26	17	9
85	10	19	28	27	9	7
86	8	9	13	19	23	28
87	13	9	19	11	25	23
88	16	21	24	16	10	13

TABLE C-9

TECHNICAL STAFF SURVEY RESPONSES TO
ITEMS 89 TO 91
(PERCENTAGES)

ITEM	RESPONSE	%
89	A. I do not know what it is.	11
	B. I know what it is but have not seen it.	2
	C. I have seen it, but have not looked through it.	2
	D. I have browsed through it quickly.	24
	E. I have read it.	22
	F. I have read it more than once.	30
	G. I have studied it thoroughly.	9
90	A. I fully understand it.	5
	B. I understand it fairly well.	42
	C. I do not understand some parts of it.	30
	D. I understand very little of it.	20
	E. I do not know anything about it.	3
91	A. I like it.	1
	B. There are some good and bad points to it, but the good probably outweigh the bad.	33
	C. There are some good and bad points to it, but the bad probably outweigh the good.	47
	D. I don't like it.	19

TABLE C-10

COMPARISON OF SUPERVISOR AND

NON-SUPERVISOR SURVEY RESPONSES

CLUSTER	SUPERVISOR (n = 26)	NON-SUPERVISORS (n = 127)	t-VALUE[a]
TRASANA	4.92[b] (.76)	4.19 (.80)	4.41
Job Satisfaction	4.84 (.84)	4.43 (.89)	2.30
Control of Work	5.06 (.93)	4.52 (1.09)	2.62
General Reaction	3.84 (.76)	3.36 (.74)	2.90
Fairness	3.56 (1.15)	2.82 (1.18)	2.98
Understanding	4.85 (.66)	3.84 (.79)	6.85
Management Tool	3.92 (.64)	3.35 (.70)	4.06
TOTAL	4.29 (.41)	3.71 (.44)	6.52

[a] *All t-values are significant beyond the .05 level (critical values: $p < .05 = 1.96$; $p < .01 = 2.58$; $p < .001 = 3.29$).*

[b] *Top number is the mean and the number in parenthesis is the standard deviation.*

TABLE C-11

TTS AND PI SAMPLE BY GRADE

GRADE	PERCENT OF TTS	PERCENT OF PI SAMPLE
CIVILIAN:		
GS-7	4.2	2.7
9	3.5	5.3
11	6.5	9.2
12	25.4	28.3
13	26.2	32.2
14	12.3	3.3
15	3.5	-
OTHER	-	1.0
MILITARY:		
Officer:	12.7	12.0
Enlisted:	5.8	6.0
TOTAL	100.1[a]	100

[a]*Does not total 100% due to rounding.*

TABLE C-12. TPMS SCORING DATA BY BRANCH

BRANCH	PROJECTS[a]	TMM	REGULAR OVERHEAD	MAN DIR OVERHEAD	JUDG AWARDS[b]	DISC AWARDS	DISC PP	TOTAL PP	PI
1	2	58.41	4.06	0.43	0	0	0	108.80	1.81
2	1	80.77	6.75	0.11	0	0	0	136.73	1.64
3	10	44.33	3.02	4.43	-2	2	2.36	72.48	1.63
4	3	31.17	0.50	0.45	0	0	0	50.92	1.63
5	8	61.99	7.80	1.01	+1	0	0	105.09	1.62
6	3	165.29	2.71	0.23	0	0	0	265.28	1.59
7	10	106.26	2.68	0.07	-1, +1	1	1.50	168.94	1.58
8	4	35.00	7.35	0.57	0	0	0	58.74	1.58
9	7	73.36	9.44	0.54	0	0	0	121.14	1.57
10	12	36.18	5.70	0.90	0	1	1.0	58.88	1.55
11	9	98.09	16.63	0.38	0	5	2.08	165.31	1.52
12	9	32.64	16.72	0.16	0	3	0.75	57.51	1.50
13	2	22.74	13.72	0.85	0	0	0	39.53	1.47
14	10	28.89	9.01	0.18	0	1	0.25	45.36	1.44
15	12	54.74	33.07	0	0	2	2.45	91.59	1.42
16	3	3.68	4.03	0.1	0	0	0	6.87	1.42
17	4	32.43	4.55	2.69	0	0	0	45.98	1.41
18	2	19.86	3.56	0	0	0	0	28.76	1.38
19	4	33.46	44.14	4.73	0	1	0.76	57.63	1.36
20	4	16.21	11.44	1.94	0	2	4.00	19.98	1.33
21	17	21.78	50.34	0	0	3	8.0	43.77	1.30
22	8	28.46	12.70	0.11	0	0	0	39.49	1.27
23	2	17.57	13.12	0.56	0	0	0	24.27	1.23
24	2	0.90	4.96	0.35	0	0	0	1.59	1.17

NOTE: These data were taken from the TPMS scoring sheets submitted by each division office. Branches are arranged from highest to lowest PI.

[a] *Includes the total number of projects for which productivity points were awarded to the branch.*

[b] *Only interbranch judgmental awards (JA) are shown. A negative JA means that the branch received up to 30 percent less PP's than earned based solely on manpower charges. A positive JA means the branch received up to 30 percent more PP's than earned based solely on manpower charges. All of these JA's were in the same division.*

TABLE C-13. TPMS SCORING DATA BY DIVISION

PROJECTS[a]	TMM	REGULAR OVERHEAD[b]		MAN DIR OVERHEAD	JUDG AWARDS[c]	DISC AWARDS	DISC PP	TOTAL PP	PI
8	280.47	11.46	(4)	1.31	0	0	0	456.56	1.60
14	213.98	26.55	(12)	5.69	-1	3	3.86	349.21	1.57
5	77.65	27.39	(35)	1.81	0	0	0	136.08	1.56
13	158.38	11.10	(7)	1.38	0	1	0.4	251.07	1.55
10	159.94	66.74	(42)	0.87	0	6	2.85	263.31	1.46
35	87.70	78.67	(90)	1.25	+1	3	9.25	152.00	1.39
17	110.01	108.16	(98)	5.80	0	2	2.45	182.81	1.35
7	48.59	31.57	(65)	2.56	0	2	4.00	72.83	1.32

NOTE: These data were taken from the TPMS scoring sheets submitted by each division office. Divisions are ordered by PI.

[a]*Includes only projects for which the division was primarily responsible.*

[b]*Number in parentheses are percent of TMM, rounded to nearest whole number.*

[c]*Only interdivision JA's are shown. One division received less PP's and another received more than earned based solely on manpower charges for one project.*

APPENDIX D

TPMS SCORING SHEETS

TPMS DIVISION SCORING SHEET

Division: ______

Discretionary Productivity Points DPP

Projects:	PROJECT SCORE (PS)	TOTAL PROJECT MAN-MONTHS EXPENDED (MM)	TOTAL PROJECT PRODUCTIVITY POINTS LESS DISCRETIONARY PRODUCTIVITY POINTS	TOTAL DIVISION MAN-MONTHS EXPENDED (DMM)	CUMULATIVE DIVISION MAN-MONTHS EXPENDED*	DIVISION PRODUCTIVITY POINTS AWARDED DUE TO DMM	DPP EXCEPTIONAL CONTRIBUTION	DPP PROJECT RESCUE	DPP UNREWARDED OUTSTANDING PERFORMANCE	DPP REDRESS OF MALDISTRIBUTION	TOTAL DIVISION PRODUCTIVITY POINTS (SUM OF ENTRIES #6, #7, ... #8)	CUMULATIVE DIVISION PRODUCTIVITY POINTS	PRODUCTIVITY INDEX (THIS PROJECT)	CUMULATIVE PI*

$$\text{Productivity Index} = \frac{\text{Sum of column \#11}}{\text{Sum of column \# 4}} = \_\_\_\_$$

1. Record overhead charge accounts as separate projects.
2. Entry in column 3 is PS x MM.
3. Entries marked with an asterisk are optional.

TPMS BRANCH SCORING SHEET

Division: ____________

Branch: ____________

Discretionary Productivity Points (DPP)

Projects:	PROJECT SCORE *	TOTAL DIVISION PRODUCTIVITY POINTS LESS DISCRETIONARY PRODUCTIVITY POINTS	TOTAL DIVISION MAN-MONTHS EXPENDED (DMM)	TOTAL BRANCH MAN-MONTHS EXPENDED (BMM)	CUMULATIVE BRANCH MAN-MONTHS EXPENDED*	BRANCH PRODUCTIVITY POINTS AWARDED DUE TO BMM	DPP EXCEPTIONAL CONTRIBUTION	DPP PROJECT RESCUE	DPP UNREWARDED OUTSTANDING PERFORMANCE	DPP REDRESS OF MALDISTRIBUTION	TOTAL BRANCH PRODUCTIVITY POINTS (SUM OF ENTRIES #6, #7, #8)	CUMULATIVE BRANCH PRODUCTIVITY* POINTS	PRODUCTIVITY INDEX* (THIS PROJECT)	CUMULATIVE* PI

$$\text{Productivity Index} = \frac{\text{Sum of column \#11}}{\text{Sum of column \# 4}} = \_\_\_\_$$

1. Record overhead charge accounts as separate projects.
2. Entries marked with an asterisk are optional.

TPMS INDIVIDUAL SCORING SHEET

Division: ______

Branch: ______

Name: ______

Project:	PROJECT SCORE*	TOTAL BRANCH PRODUCTIVITY POINTS LESS DISCRETIONARY PRODUCTIVITY POINTS	TOTAL BRANCH MAN-MONTHS EXPENDED (BMM)	TOTAL INDIVIDUAL MAN-MONTHS EXPENDED (IMM)	CUMULATIVE INDIVIDUAL MAN-MONTHS* EXPENDED	INDIVIDUAL PRODUCTIVITY POINTS AWARDED DUE TO IMM	Discretionary Productivity Points (DPP): DPP EXCEPTIONAL CONTRIBUTION	DPP PROJECT RESCUE	DPP UNREWARDED OUTSTANDING PERFORMANCE	DPP REDRESS OF MALDISTRIBUTION	TOTAL INDIVIDUAL PRODUCTIVITY POINTS (SUM OF ENTRIES #6, #7, #8, #9, and #10)	CUMULATIVE INDIVIDUAL PRODUCTIVITY* POINTS	PRODUCTIVITY INDEX (PI)* (THIS PROJECT)	CUMULATIVE* PI

$$\text{Productivity Index} = \frac{\text{Sum of column \#11}}{\text{Sum of column \# 4}} = \_\_\_\_\_$$

(overhead terms may be deleted)

1. Record overhead charge accounts as separate projects.
2. Entries marked with an asterisk are optional.

APPENDIX E

MATHEMATICAL ANALYSIS OF THE TPMS SCORING ALGORITHM

APPENDIX E

MATHEMATICAL ANALYSIS OF THE TPMS SCORING ALGORITHM

1. GENERAL

This appendix provides the detailed mathematical analysis of the TPMS scoring algorithm.

2. THE TPMS MEASUREMENT FUNCTION

The assignment of the productivity index (PI) by the measurement function can be conveniently segmented into a sequence of assignments by component functions. First, a project scoring function yields a project score (S) from relevant variables. Next, using the concept of productivity points (PP), this score is apportioned among the units and individuals who contributed to the project in relation to manpower expended and to the quality of the contribution. Finally, the PI for the individual or the unit is given as an average of project PP earned, weighted by the manpower expended on the projects.

a. Project Scoring Function

The S is the product of five project variables: manpower utilization efficiency (R), computer utilization efficiency (C), timeliness (T), quality (Q), and priority (P). Thus

$$S = R \times C \times T \times Q \times P$$

Each of these five factors is discussed separately.

(1) Priority. Originally the value of P was to vary from 1.0 to 1.2 with the value of 1.2 assigned to external projects. For internal projects the value of P was to be determined by the PMB. Since all projects were assigned a P value of 1.2 during the first scoring period, this variable will be held constant in the subsequent analysis.

(2) Quality. The value of Q is assigned by the PRB upon review of the project final report. Values of Q may vary from 1.00 to 1.50.

(3) Manpower Utilization Efficiency. The value of R is determined from the planned manpower usage for the project (R_P) and the actual manpower used on the project (R_A) as follows:

$$R = \begin{cases} 0.7 \text{ when } R_P/R_A < 0.5 \\ 1.0 \text{ when } R_P/R_A = 1.0 \\ 1.2 \text{ when } R_P/R_A > 2.0 \end{cases}$$

Intermediate values of R scale up linearly between 1.0 and 1.2 and scale down linearly between 1.0 and 0.7. The expression for R as a piecewise linear function of R_P/R_A is then

$$R = \begin{cases} 0.7 & \text{when } R_P/R_A < 0.5 \\ 0.6\left(R_P/R_A\right) + 0.4 & \text{when } 0.5 \leq R_P/R_A < 1.0 \\ 0.2\left(R_P/R_A\right) + 0.8 & \text{when } 1.0 \leq R_P/R_A < 2.0 \\ 1.2 & \text{when } 2.0 \leq R_P/R_A \end{cases}$$

The graph of this function is shown in figure E-1.

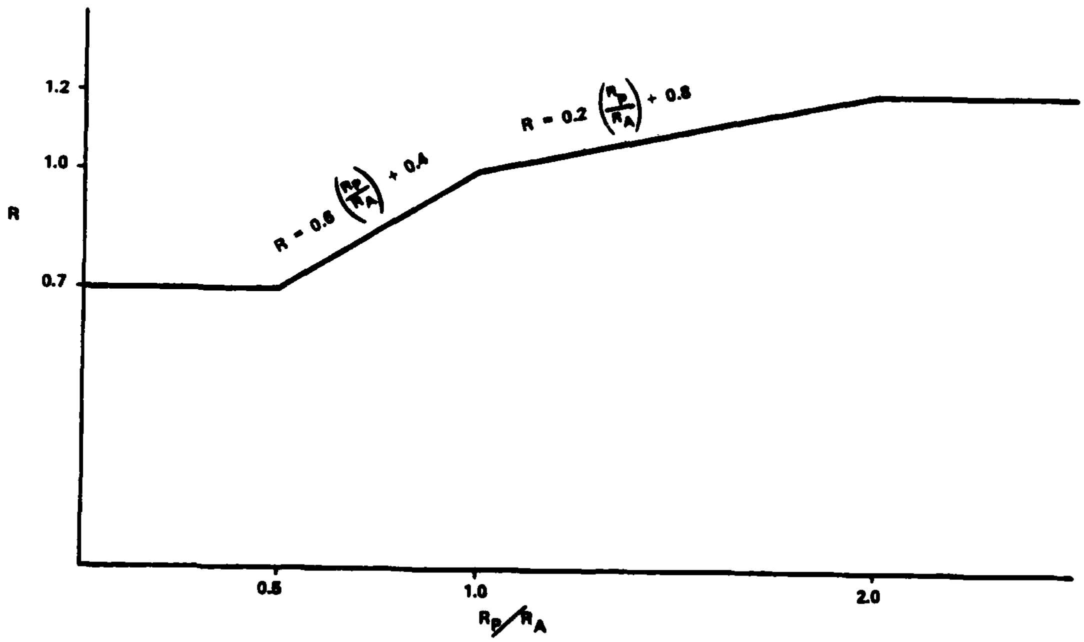


Figure E-1. Manpower Utilization Efficiency (R)

(4) Computer utilization efficiency. The value of C is determined from the planned computer usage for the project (C_P) and the actual computer usage (C_A) as follows:

$$C = \begin{cases} 0.9 \text{ when } C_P/C_A < 0.5 \\ 1.0 \text{ when } C_P/C_A = 1.0 \\ 1.1 \text{ when } C_P/C_A > 2.0 \end{cases}$$

Intermediate values of C scale up and down linearly from C = 1.0. The expression for C as a piecewise linear function of C_P/C_A is:

$$C = \begin{cases} 0.9 & \text{when } C_P/C_A < 0.5 \\ 0.2\,(C_P/C_A) + 0.8 & \text{when } 0.5 \leq C_P/C_A < 1.0 \\ 0.1\,(C_P/C_A) + 0.9 & \text{when } 1.0 \leq C_P/C_A < 2.0 \\ 1.1 & \text{when } 2.0 \leq C_P/C_A \end{cases}$$

Figure E-2 is the graph of this function. During the course of the first scoring period the measure of computer usage was changed. Because of this, 51 of the 61 projects scored in the first period were administratively assigned a C value of 1.

(5) Timeliness. The value of T is a function of the planned completion date and the actual completion date. The value may vary between a maximum value of 1.1 for early completion and a minimum value of 0.8 for late completion. The functional relationship yielding T is project dependent and is determined by the project study manager with the approval of the PMB. Figure E-3 gives the timeliness value as a function of time for one project.

(6) Project score values. Values for project scores can vary from 0.504 to 2.6136.[1] A value of 1.56 might be viewed as a "base" score - representing a project with minimally acceptable quality (B or Q = 1.3), exactly on time (T = 1.0), and using exactly planned manpower (R = 1.0) and computer usages (C = 1.0). The minimum score for a project of minimally acceptable quality is 0.7862. Table E-1 shows maximum, minimum, "base", and minimally acceptable project scores with C varying in the first case, and held constant at 1.0 in the second case. The project score has a wider range of possible values when C is not held constant.

---

[1] *All projects scored in the first period have P = 1.2. When calculated with all variables equal to the lowest value except for P = 1.2, the lowest score is 0.6048.*

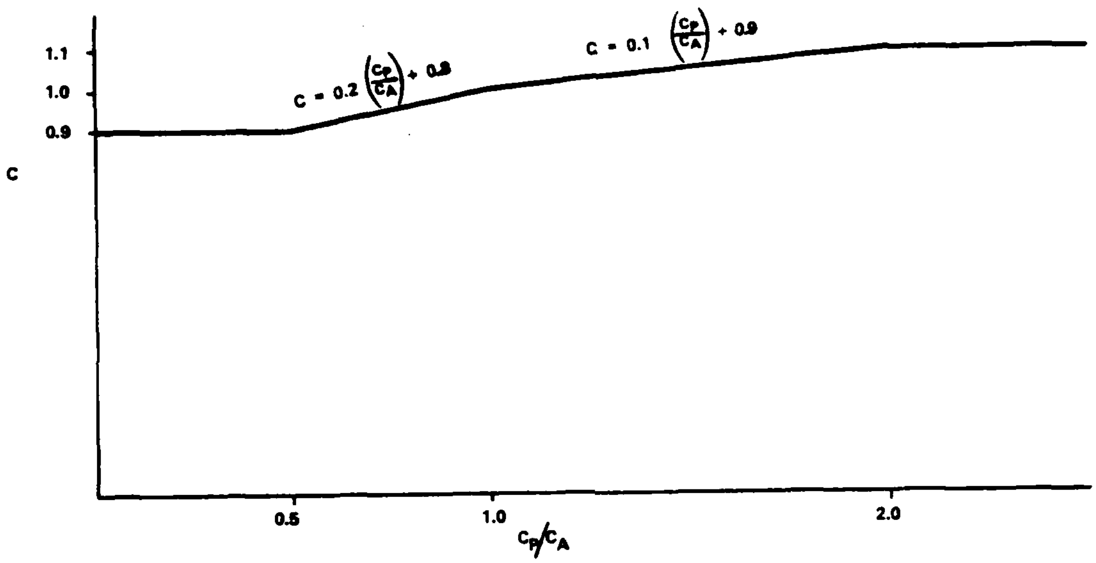


Figure E-2. Computer Utilization Efficiency (C)

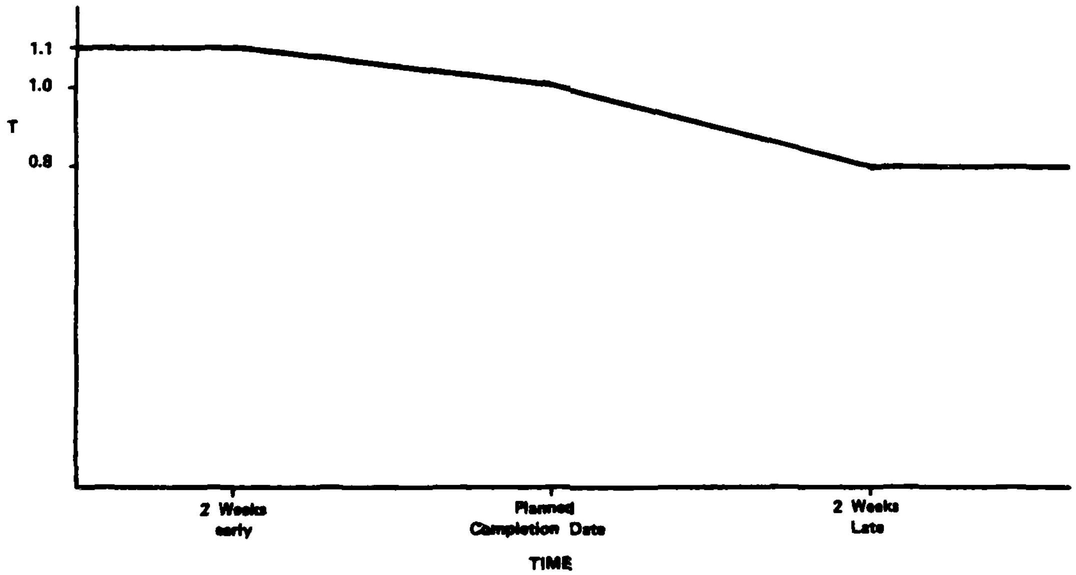


Figure E-3. Timeliness (T)

TABLE E-1

PROJECT SCORES FOR REPRESENTATIVE INPUTS

	Variable Value of C					
Case	P	Q	R	C	T	S
Maximum	1.2	1.5	1.2	1.1	1.1	[illegible]
Minimum	1.2	1.0	0.7	0.9	0.8	0.6048
Base	1.2	1.3	1.0	1.0	1.0	1.56
Minimally Acceptable	1.2	1.3	0.7	0.9	0.8	0.78624
	Constant value of C					
Case	P	Q	R	C	T	S
Maximum	1.2	1.5	1.2	1	1.1	2.592
Minimum	1.2	1.0	0.7	1	0.8	0.672
Base	1.2	1.3	1.0	1	1.0	1.56
Minimally Acceptable	1.2	1.3	0.7	1	0.8	0.8736

b. Productivity Point Distribution

Productivity points are generated for a project by multiplying the project score by the actual manpower usage (in man-months).[2]

$$PP = S \times R_A$$

These PP are then distributed among the units and individuals who charged time to the project. Basically an individual or unit is awarded PP in relation to the manpower charged to the project. However, the quality of an individual's or unit's contribution can result in a plus or minus 30 percent deviation from

[2]Overhead can be treated as a project with a score of 1 for general overhead or 2 for management directed overhead.

this straight charge to manpower award. In addition, the distribution of PP is done in a hierarchical manner which makes deviations beyond 30 percent possible in some cases. Points are distributed first to divisions, then to branches within a division, and finally to individuals within a branch. At any of the 3 levels in the distribution a plus or minus 30 percent deviation from a straight charge to manpower award is possible.

(1) Division. If division j charges M_j man-months to a project it earns

$$PP \times (M_j/R_A) \times d_j$$

or

$$S \times M_j \times d_j \text{ productivity points}$$

$$\text{where } 0.7 \leq d_j \leq 1.3$$

The d_j is the division judgmental factor. If the award is strictly according to manpower charged, $d_j = 1$. Since all manpower charged originates from some division and all PP must be awarded:

$$R_A = \sum_{\substack{\text{over} \\ \text{contributing divisions}}} M_j$$

and

$$R_A = \sum_{\substack{\text{over} \\ \text{contributing divisions}}} (M_j \times d_j)$$

(2) Branch.[3] The $S \times M_j \times d_j$ productivity points distributed to division j are next awarded to branches of this division which contributed to the project. If branch k of division j charged M_{jk} man-months to the project it earns

$$S \times M_j \times d_j \times (M_{jk}/M_j) \times b_k$$

or

$$S \times M_{jk} \times d_j \times b_k \text{ productivity points}$$

$$\text{where } 0.7 \leq b_k \leq 1.3$$

---

[3] *In the hierachical distribution of PP, persons at the division level (not assigned to a branch) are award PP at this level. For this discussion, division staff personnel are treated the same as a branch.*

The b_k is the branch judgmental factor. If the award is strictly according to manpower charged, $b_k = 1$. Since all divisional manpower charged originates from some branch and all PP must be awarded:

$$M_j = \sum_{\substack{\text{over contributing} \\ \text{branches of division } j}} M_{jk}$$

and

$$M_j \times d_j = \sum_{\substack{\text{over contributing} \\ \text{branches of division } j}} (M_{jk} \times d_j \times b_k)$$

(3) Individual. The $S \times M_{jk} \times d_j \times b_k$ productivity points distributed to branch k are next distributed to the individuals within this branch who contributed to the project. If individual l of branch k of division j charges M_{jkl} man-months to the project, he earns

$$S \times M_{jk} \times d_j \times b_k \times (M_{jkl}/M_{jk}) \times p_l$$

or

$$S \times M_{jkl} \times d_j \times b_k \times p_l \text{ productivity points}$$

where $0.7 \leq p_l \leq 1.3$

The p_l is an individual judgmental factor. If the award is strictly according to manpower charged, $p_l = 1$. Since all branch manpower originates from some individual and all PP must be distributed:

$$M_{jk} = \sum_{\substack{\text{contributing} \\ \text{individuals} \\ \text{of branch } k}} M_{jkl}$$

$$M_{jk} \times d_j \times b_k = \sum_{\substack{\text{contributing} \\ \text{individuals} \\ \text{of branch } k}} (M_{jkl} \times d_j \times b_k \times p_l)$$

(4) Discretionary Productivity Points (DPP). In addition to the productivity points generated by the project score and manpower expended, there is another class of points - discretionary productivity points. These points are awarded to units and individuals whose contributions to a project are not adequately rewarded by the standard point distribution process. TRASANA Memo 5-4 describes DPP and the constraints on their awarding.

c. Assigning the Productivity Index (PI)

The PI is an average of project PP weighted by the manpower expended on the projects. This assignment is described in detail for divisions, branches, and individuals.

(1) Division. If division j contributes to n projects (including overhead) for i = 1, 2, 3,...., n with scores S_i, expending M_{ij} man-months, earning DPP_{ij}[4] then

$$PI = \frac{\sum_i (S_i \times M_{ij} \times d_{ij}) + DPP_{ij}}{\sum_i M_{ij}}$$

$$= \sum_i S_i \times d_{ij} \times \left(\frac{M_{ij}}{\sum_i M_{ij}} \right) + \frac{\sum_i DPP_{ij}}{\sum_i M_{ij}}$$

(2) Branch. If branch k of division j contributes to n projects (including overhead) for i = 1, 2,..., n with scores S_i, expending M_{ijk} man-months, earning DPP_{ijk}[5] then

$$PI = \frac{\sum_i (S_i \times M_{ijk} \times d_{ij} \times b_{ik}) + DPP_{ijk}}{\sum_i M_{ijk}}$$

$$= \sum_i S_i \times d_{ij} \times b_{ik} \times \left(\frac{M_{ijk}}{\sum_i M_{ijk}} \right) + \frac{\sum_i DPP_{ijk}}{\sum_i M_{ijk}}$$

---

[4] *At the division level only DPP for exceptional contribution and project rescue are counted.*

[5] *At the branch level only DPP for exceptional contribution and project rescue are counted.*

(3) Individual.[6] If individual l of branch k and division j contributes to n projects (now overhead may be deleted) i = 1,2,..., n with scores S_i, expending M_{ijkl} man-months, earning DPP_{ijkl} then

$$PI = \frac{\sum_i (S_i \times M_{ijkl} \times d_{ij} \times b_{ik} \times p_{il}) + DPP_{ijkl}}{\sum_i M_{ijkl}}$$

$$= \sum_i \left(S_i \times M_{ijkl} \times d_{ij} \times b_{ik} \times p_{il} \times \frac{M_{ijkl}}{\sum_i M_{ijkl}}\right) + \frac{\sum_i DPP_{ijkl}}{\sum_i M_{ijkl}}$$

3. MEASUREMENT ALGORITHM ERROR ANALYSIS

Any measurement contains some error. The usefulness of the measurement can only be fully realized when the magnitude of the possible error is clearly specified. This analysis provides approximate bounds for the numerical error associated with the project score and the PI. Numerical error in the computed values of a function is dependent upon the algorithm chosen for the computation. Once a correct algorithm is chosen, numerical error may be caused by (1) error in input values, (2) algorithmic error (i.e. round-off error), and (3) mistakes (i.e. arithmetical). The project scoring sheets indicated the use of fairly consistent procedures for computing project scores from the input variables. The analysis of numerical error in project scores assumes these procedures.[7] Standard forms are used to compute the PI given the project scores. The computational procedure suggested by these forms will also be assumed for the analysis of the error in the PI.

a. Project Score Error

First the error in each of the four factors is discussed, and then total error is described.

(1) Quality. The value of Q is given by the PRB to two decimal places, resulting in an absolute error of no more than 0.005. Because the value of Q is near 1, the maximum absolute relative error can be taken to be this same value (0.005).

---

[6] *The PI for division level staff is computed by a different formula. Refer to TRASANA Memo 5-4.*

[7] *See Conte and de Boor (1980) for a discussion of numerical error analysis.*

(2) Manpower utilization efficiency.

(a) Project scoring sheets indicate that the formula for computing R is

$$R = \begin{cases} 0.7 & \text{when } R_P/R_A < 0.5 \\ 1 - 0.3\left(\dfrac{R_A - R_P}{R_P}\right) & \text{when } 0.5 \leq R_P/R_A < 1.0 \\ 1 + 0.2\left(\dfrac{R_P - R_A}{0.5\ R_P}\right) & \text{when } 1.0 \leq R_P/R_A < 2.0 \\ 1.2 & \text{when } 2.0 \leq R_P/R_A \end{cases}$$

This interpolation formula is inconsistent with the previous definition of R in terms of R_P/R_A. In fact, the above formula does not give R as a piecewise linear function of R_P/R_A, but rather

$$R = \begin{cases} 0.7 & \text{when } R_P/R_A < 0.5 \\ 1.3 - 0.3\left(R_P/R_A\right)^{-1} & \text{when } 0.5 \leq R_P/R_A < 1.0 \\ 1.4 - 0.4\left(R_P/R_A\right)^{-1} & \text{when } 1.0 \leq R_P/R_A < 2.0 \\ 1.2 & \text{when } 2.0 \leq R_P/R_A \end{cases}$$

(b) Assuming error free input values, the use of the interpolation formula rather than the formula of paragraph 2a(3) could cause the value of R to be in error by as much as 0.05. Since the value of R is rounded to two decimal places before recording on the project scoring sheet and inclusion into the multiplicative formula for S, the absolute maximum error could be as much as 0.055. Since R is close to 1, 0.055 would also be an approximation of the maximum absolute relative error.

(c) A more useful analysis of the error in R would assume the use of the formula of paragraph 2a(3) for the computation. The value of R_P could be assumed to be exact, so that input error is entirely the result of R_A. The value of R_A (in man-months) to two decimal places is provided by the TRASANA Manpower Utilization Reporting System (TMURS).[8] The maximum absolute input

[8] *The most recent version of TMURS reports man-months to 3 decimal places.*

error of 0.005 in R_A yields an approximate maximum absolute error of 0.001 in R. Thus 0.001 approximates the maximum absolute relative error in R, since R is near 1.

(d) The TMURS manpower accounting system collects, compiles, and reports manpower charges for all types of activities. Manpower charges are recorded in man-hours and reported in both man-hours and man-months. However, the conversion from man-hours to man-months varies by month. Thus, man-month as reported by TMURS is not a fixed unit of measure: one person working [illegible] of February, 160 man-hours, and one person working each work[illegible] day of October, [illegible] man-hours, each are reported as having worked 1 man-month. If this fact is not taken into account when R_P is determined, manpower over-runs or under-runs [illegible] result. For computing R, a fixed unit should be used for measuring both R_P and R_A. This could be man-hours, man-weeks (40 man-hours), or man-months with one man-month equal to a standard number of hours, such as 168 man-hours.

(3) Computer utilization efficiency. During the first scoring period most projects were scored with an exact C value of 1. Since this will not be the case for subsequent scoring periods, the numerical error due to variable C is described. This description assumes the use of the formula of paragraph 2a(4) and that the value of C_P is exact. Thus all input error is a consequence of C_A. The value of C_A is measured in Standard Units of Processing (SUP) which consists of the sum of Central Processing Unit (CPU) time, Input/Output (I/O) time, and Control Card/Executive Request (CC/ER) time in hours. Suppose that the value of C_A is reported in SUP hours to the nearest hundredth (3.6 seconds). This error in C_A results in an approximate maximum absolute error of 0.001 in C. Because C is near 1, 0.001 approximates the maximum absolute relative error in C.

(4) Timeliness. The timeliness factor is a function of the deviation of the actual delivery date of the completed project report from the planned delivery date. This functional relationship is project dependent. A detailed error analysis would require this function in addition to knowledge of the accuracy with which the deviation in dates is measured (e.g. nearest day, nearest week, etc.). If the deviation can be assumed to be exact (in days, say) the numerical error in T is solely a consequence of the current practice of rounding T to two decimal places before its incorporation into the formula for S. The maximum absolute relative error is then 0.005.

(5) Project Score. The relative error in S is simply the sum of the relative errors of its factors. Since S is near 1, relative error approximates actual error. For the first scoring period (using the interpolation formula for R) the project score has maximum absolute (relative) error of approximately 0.071. This result assumes exact arithmetic in the computations except for the intermediate roundings of R, C, and T. A final rounding to two decimal places results in an approximate error of 0.076 in the project score. A more useful result would assume that the formula of paragraph 2a(3) is used to compute R. If the value of T is assumed to be exact and all arithmetical computations are exact, the maximum absolute (relative) error in S is approximately 0.007. If intermediate roundings of R, C, and T, a final rounding of S are done to two decimal places, the maximum numerical error in S is approximately 0.027. Bounds for the numerical error in the project score are

summarized in table E-2. The results presented there assume exact arithmetic except where otherwise noted. A numerical example is given in table E-3 to illustrate the efficiency of the bound (0.027) from the last row and last column of table E-2. In table E-3 the value of S given in the first column and the rounded value of S given in the last column differ by almost 0.026, thus demonstrating the efficiency of the 0.027 bound. A reported project score could then be in error by as much as plus or minus 0.027. For example, if a reported project score is 1.560 (base case) the true value could vary anywhere from 1.587 to 1.533.

TABLE E-2

SUMMARY OF NUMERICAL ERRORS (APPROXIMATE MAXIMUM ABSOLUTE ERROR)

VARIABLE	Interpolation Formula for R		Formula of 2a(3) for R	
	NO INTERMEDIATE ROUNDING	INTERMEDIATE ROUNDING	NO INTERMEDIATE ROUNDING	INTERMEDIATE ROUNDING
O	.005	.005	.005	.005
R	.05	.055	.001	.006
C	.001	.006	.001	.006
T	NA	.005	NA	.005
S	.056	.071	.007	.022
S ROUNDED TO TWO DECIMAL PLACES	.061	.076	.012	.027

b. Productivity Index Error

Following the formulas of 2c the PI may be computed by summing PP and manpower expended and dividing the two. Using exact arithmetic, the maximum absolute error in a computed PI is bounded by the maximum absolute error in the project scores plus a much smaller term involving the number of projects, total manpower, and total PP. In fact, the maximum absolute error (0.027) in the project score alone is a realistic approximation of the numerical error in the PI. Thus a reported PI could be in error by as much as plus or minus 0.027. In the first scoring year there was only a difference of 0.03 between the lowest PI of the top nine scoring individuals and the PI of the next highest scorer. This difference could in fact then be due almost entirely to numerical error.

TABLE F-3

EXAMPLE OF NUMERICAL ERROR IN COMPUTING A PROJECT SCORE

$R_P = C_P = 3$. All computations done with ten significant figures

VARIABLE	EXACT INPUTS R&C GIVEN TO TEN SIGNIFICANT FIGURES	APPROXIMATE INPUTS R&C ROUNDED TO TWO DECIMAL PLACES
Q	1.[illegible]5	1.[illegible]
R_A	3.242	[illegible].24
P	0.954270[illegible]	0.9[illegible]
C_A	3.244	3.24
C	0.9849568434[illegible]	0.99
T	1.005	1.01
S	1.224045389[b]	1.247875200
S ROUNDED TO TWO DECIMAL PLACES	1.22	1.25

[a]ERROR < 3.0×10^{-10}
[b]ERROR < 1.0×10^{-9}

4. RELATIONSHIPS AMONG THE PROJECT SCORE VARIABLES

A traditional measure of productivity is value of output divided by value of input. The project scoring function which is a global measure of a project team's productivity is a variation of this traditional measure. If f and g denote the piecewise linear functions of paragraphs 2a(3) and (4) respectively, then

$$S = P \times Q \times T \times f\left(R_P/R_A\right) \times g\left(C_P/C_A\right)$$

In a sense, R_P and C_P represent the output values of a project (as determined by the PMB) and R_A and C_A represent input values. Rather than use ratios directly, they are adjusted by f and g to take into account the relative costs of these two resources. The Q and T (and also P) factors might be viewed as "value multipliers." Timeliness or quality increase the worth of a project, or perhaps more realistically, the lack of timeliness or quality decrease its worth.

a. Schemes for Assembling Factors

The method for assembling the individual factors into a project score, the project scoring function, is a combined multiplicative and conjunctive scheme. To contrast alternate scoring schemes, suppose that the project score S is a function of Q and R alone.

(1) An additive or compensatory scheme would compute S as $S = Q + R$. Using $Q = 1.3$ and $R = 1.0$ as base values for a B or "good" project, the category of B or better projects is defined by $S \geq 2.3$. The compensatory property of this scheme is illustrated by the fact that a project with $Q = 1.2$ and $R = 1.1$ and a project with $Q = 1.4$ and $R = 0.9$ both score 2.3.

(2) A conjunctive scheme would define a B or better project as one with $Q \geq 1.3$ and $R \geq 1.0$. Now a low score on one factor can not be compensated for by a high score on another factor.

(3) A strictly multiplicative scheme for S would compute S by $S = Q \times R$. A project graded B or better would be defined as one with $S \geq 1.3 = (1.3)\ (1.0)$. To some extent this scoring function is also compensatory. For example, a project with $Q = 1.35$ and $R = 0.962$ and a project with $Q = 1.2$ and $R = 1.0833$ both have $S = 1.3$. Thus a low value on one factor can be compensated for by a higher value on another.

(4) Although the TPMS scoring function is multiplicative, to avoid the compensatory property (at least for Q) a minimum acceptable value of $Q = 1.3$ must be achieved. In theory a project scoring less than 1.3 on Q would be required to expend more manpower to upgrade quality to at least 1.3. In computing S the original Q value is used with the inflated R value. In practice, because some R scores are so high, low scores for Q are in effect compensated for. In the first scoring period for example, one project had $Q = 1.2$ and $R = 1.122$ for $S = 1.3464$.[9] Another project had $Q = 1.35$ and $R = .998$ for $S = 1.3474$. Thus it appears that a low score on Q is compensated for by a high score on R and *vice versa*.

(5) Figures E-4 to E-7 illustrate the different scoring schemes just discussed. The shaded area in each corresponds to scores of B or better.

b. Distinct Resource Efficiency Factors

Two resource efficiency factors are used to compute the project score. Two difficulties arise as a consequence of these factors.

(1) Scale factors. After the ratio of planned resource to actual resource is computed, it is scaled by a piecewise linear function before it is factored into the project score. These scale functions (f and g in the notation of the previous section) are the source of one difficulty.

---

[9] *For illustration only,* $S = Q \times R$.

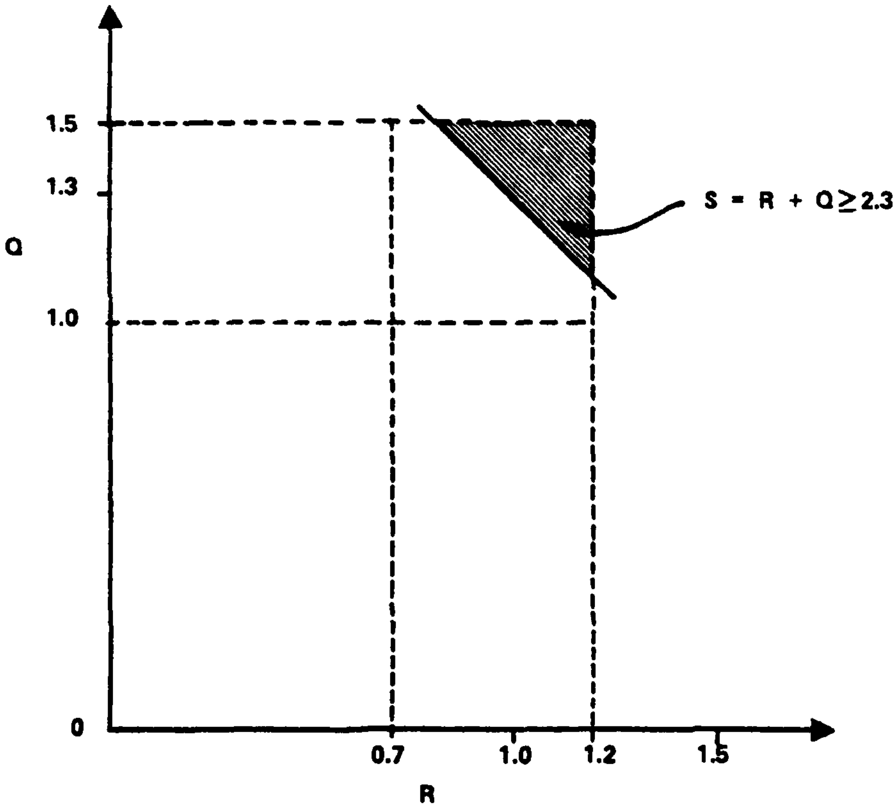


Figure E-4

Compensatory or Linear Scheme

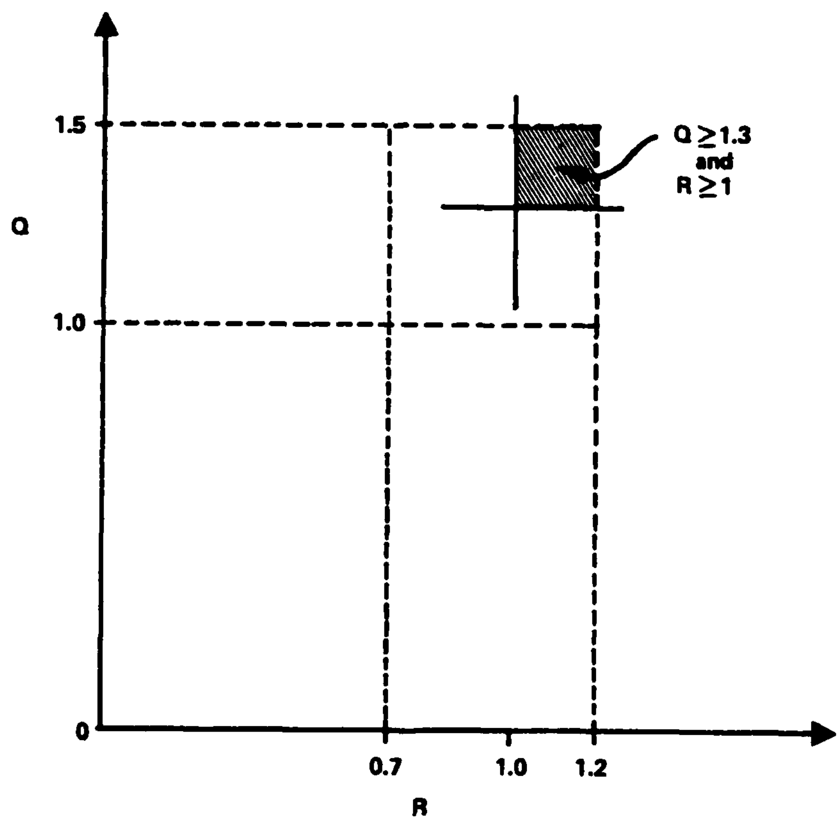


Figure E-5

Conjunctive Scheme

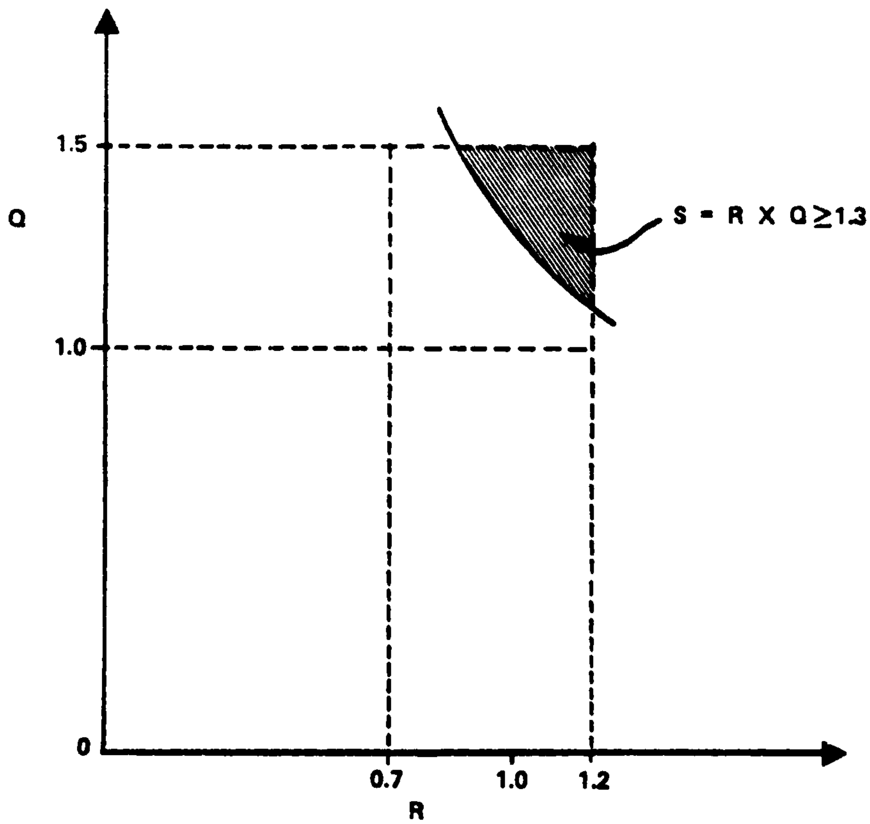


Figure E-6

Multiplicative Scheme

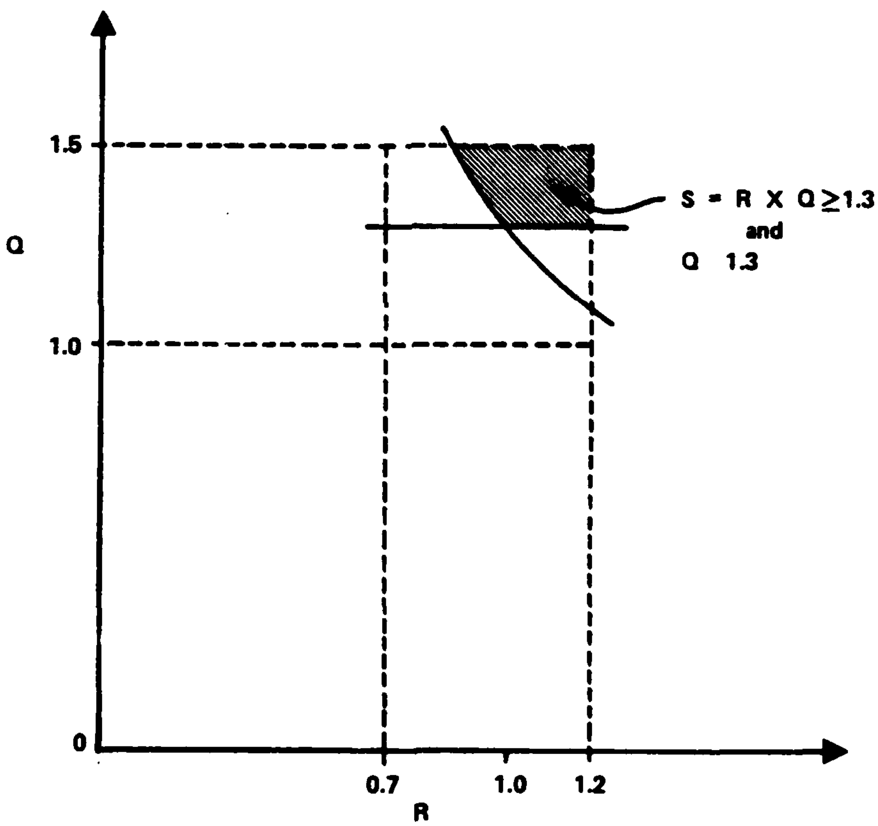


Figure E-7

Multiplicative and Conjunctive

(a) If a job requires \$10 in manpower usage and it is accomplished with \$5 worth, the ratio efficiency is 2. The TPMS scale function would assign a manpower efficiency factor of 1.2. If a job requires \$10 in computer usage and it is done with \$5 worth, the ratio efficiency is also 2. The TPMS scale function would assign a computer efficiency factor of 1.1. The rationale for different scale functions is that the absolute computer costs are small compared to absolute manpower costs, and that with the newly acquired computing system the activity has almost unlimited computing capacity. However, \$1 of resources is still \$1 of resources whether it be manpower or computer usage. Since the TPMS utilization factors are computed from ratios of planned resources to actual resources, the TPMS project scoring function seems to imply that \$1 of manpower is not equivalent to \$1 of computer usage.

(b) If the activity believes that it has surplus computer capacity and that it should encourage more use of the computer, it might be best to eliminate the computer factor from the scoring function altogether.

(2) Distinct Factors. Another rationale for having different scale functions is that it might encourage trading relatively cheap computer use for more expensive manpower. However, because the project score is a product of distinct factors, in general the optimal procedure is to maximize each factor independently of the other. In fact it is not hard to construct examples where cheap computer usage might be replaced by expensive manpower in order to increase the project score. For example, suppose a project is initiated with $R_P = 20$ and $C_P = 2$. Suppose that well into the study, the study manager is confronted with the possibility of completing the project with $R_A = 15$ and $C_A = 2$. In this case $R \times C = 1.07$. Having already used 1 SUP, he does have the option of doing the 1 SUP of computer work remaining with 5 man-months of manpower. If he chooses this option the project will be completed with $R_A = 20$ and $C_A = 1$. In this case $R \times C = 1.1$. If all other factors are the same, replacing computer usage by manpower would result in a higher project score. If it is desired that computer use replace manpower when it is cost effective, realistic costs for each resource should be given and the project score function should use a single factor which includes both manpower and computer use in relation to their costs.

c. The Differential of S

The differential (dS) of S is useful in assessing effects of changes in the values of the variables on the value of S. In general, the differential of a function can be used to approximate changes in function values given changes in its independent variables. In this case the differential is

$$dS = P\,[TRC\,\Delta O + ORC\,\Delta T + OTC\,\Delta R + OTR\,\Delta C]$$

Because O is scaled with larger values than T, R, or C; small changes ($\Delta C, \Delta T, \Delta R$) in these variables will have a greater effect on the value of S than will

a small change (ΔQ) in Q.[10] Table E-4 illustrates the effect of incrementing the variables by 0.1. Table E-5 does the same for an increment of 0.005. Notice that a smaller value of S results from incrementing Q. Scaling Q with larger values has an effect of emphasizing the other variables in the formula for S.

TABLE E-4

THE EFFECT ON THE BASE SCORE OF INCREMENTING

VARIABLES BY 0.1

CASE	Q	T	R	C	S
BASE	1.3	1	1	1	1.560
A QUALITY	1.4	1	1	1	1.680
BEST TIME	1.3	1.1	1	1	1.716
$R_P=1.5\ R_A$	1.3	1	1.1	1	1.716
$C_P=2C_A$	1.3	1	1	1.1	1.716

TABLE E-5

THE EFFECT ON THE BASE SCORE OF INCREMENTING

VARIABLES BY 0.05

CASE	Q	T	R	C	S
BASE	1.3	1	1	1	1.560
B+ QUALITY	1.35	1	1	1	1.620
GOOD TIME	1.3	1.05	1	1	1.638
$R_P=1.25\ R_A$	1.3	1	1.05	1	1.638
$C_P=1.5C_A$	1.3	1	1	1.05	1.638

[10] *In the first scoring period the smallest attained value of Q (1.15) was greater than maximum possible values for T (1.1) or C (1.1). Also for all except 3 projects the value of Q was greater than or equal to the maximum value of R.*

5. JUDGMENTAL PRODUCTIVITY POINTS

The project score is a global measure of the project team's relative productivity. The mechanism for awarding greater or fewer PP than that alloted by a straight charge to manpower award allows a manager to differentiate among the quality of individuals' or units' contributions.

a. In the hierarchical process of awarding PP a judgmental factor is included at each step of the process (division, branch, individual). This judgmental factor can range from 0.7 to 1.3 depending on the quality of contribution of that unit. The factor equals 1 if the point distribution is awarded strictly according to manpower usage. At the extremes, an individual's straight charge to manpower award could be increased by as much as a factor of 2.197 (1.3^3) or decreased by as much as a factor of 0.343 (0.7^3).

b. The hierarchical distribution of judgmental PP and compounding effect of multiplying by the judgmental factors might result in unintended results. Consider the example illustrated in table E-6. Suppose it is the judgment of the study manager and division chief that persons 2 and 3 did equally badly and should receive no more than the absolute minimum award. Person 1 carried the project and should receive the maximum possible award. As shown in the table, since persons 2 and 3 belong to different branches, person 2 automatically receives more PP and a higher PI.

TABLE E-6

AN EXAMPLE OF AWARDING JUDGMENTAL POINTS

$S = 1.5$
$R_A = 12$
$PP = 18$

	MM	PP	PI		MM	PP	PI
BRANCH 1	8	13.8	1.7	BRANCH 2	4	(1/3)(18)(.7) 4.2	1.05
Person 1 (max award)	4	(1/2)(13.8)(1.3) 8.97	2.24	Person 3 (min award)	4	4.2	1.05
Person 2 (min award)	4	(1/2)(13.8)(.7) 4.83	1.21				

6. WEIGHTED AVERAGE

The PI is simply a weighted average over a scoring period. Specifically it is an average of project scores weighted by their manpower requirements. Here overhead is considered as a project with a score of 1. Because however, a non-overhead project could carry over from one period to the next while overhead would not, spurious results can be obtained. The data in table E-7 illustrates just this. The branches seem to be equally productive during the two years, however this is not reflected in their PI's.

TABLE E-7

AN EXAMPLE OF COMPUTING BRANCH PI'S AS WEIGHTED AVERAGES

	1ST PERIOD	2ND PERIOD
BRANCH A	1 TMM Overhead 2 TMM on Project with S = 1.5 9 TMM carried over............ total TMM = 3 total PP = 4 PI = 4/3 = 1.333	1 TMM OVERHEAD 11 TMM on project with S = 1.5 total TMM = 21 total PP = 31 PI = 31/21 = 1.476
BRANCH B	1 TMM Overhead 11 TMM on project with S = 1.5 total TMM = 12 total PP = 17.5 $PI = \frac{17.5}{12} = 1.458$	1 TMM Overhead 11 TMM on project with S = 1.5 total TMM = 12 total PP = 17.5 $PI = \frac{17.5}{12} = 1.458$

APPENDIX F

AN ALTERNATIVE TPMS SCORING PROCEDURE

APPENDIX F

AN ALTERNATIVE TPMS SCORING PROCEDURE

1. GENERAL

The purpose of this study was to evaluate the first year functioning of TPMS; however, possible system improvements which followed directly from the evaluation were to be offered. This appendix presents a conceptual scoring alternative which has as its basis improving system functioning in light of the original system design attributes. This alternative is simple, would improve the system flexibility, and should considerably lessen resistance to TPMS because it specifically addresses those issues most frequently cited as problems with the current system.

2. SCORING FACTORS

The alternative procedure retains the manpower utilization efficiency factor (R), the quality factor (Q), and the timeliness factor (T). A distinct computer utilization efficiency factor is not used. As long as the computer remains a somewhat unlimited resource with fixed cost, and manpower costs account for approximately 80 percent of TRASANA resources, a separate computer factor is not seen as cost effective. If the magnitude of computer use and costs should change in the future, computer usage could be incorporated into the scoring procedure. If this is done however, a more appropriate method might be to have a combined resource efficiency factor which incorporates manpower and computer usage. This combined resource efficiency factor could then be weighted to reflect the relative cost of the two resources. Each of the three factors is discussed separately.

a. Manpower Utilization Efficiency Factor (R)

Figure F-1 is the graph of a generic efficiency factor R_P/R_A as a function of R_A/R_P. Figure F-2 is the current efficiency factor R, also as a function of R_A/R_P. The alternative suggested for scoring R is in Figure F-3. The vertical scaling is just ordinal, i.e., simply three categories of efficiency. A project with actual manpower within plus or minus 10 percent of planned manpower would be placed in the medium category. When actual manpower is 10 percent over planned manpower, it is placed in the low category. Projects with actual manpower of less than 90 percent of planned manpower are placed in the high category. In addition, a division level check on R should be implemented. For example a division might be required to have a given distribution of projects among the three categories, or bonus points could be awarded for achieving a relatively symmetrical distribution with over half of the projects in the middle category. Another possibility is to give rewards or bonus points to divisions whose average R_P/R_A is near 1.

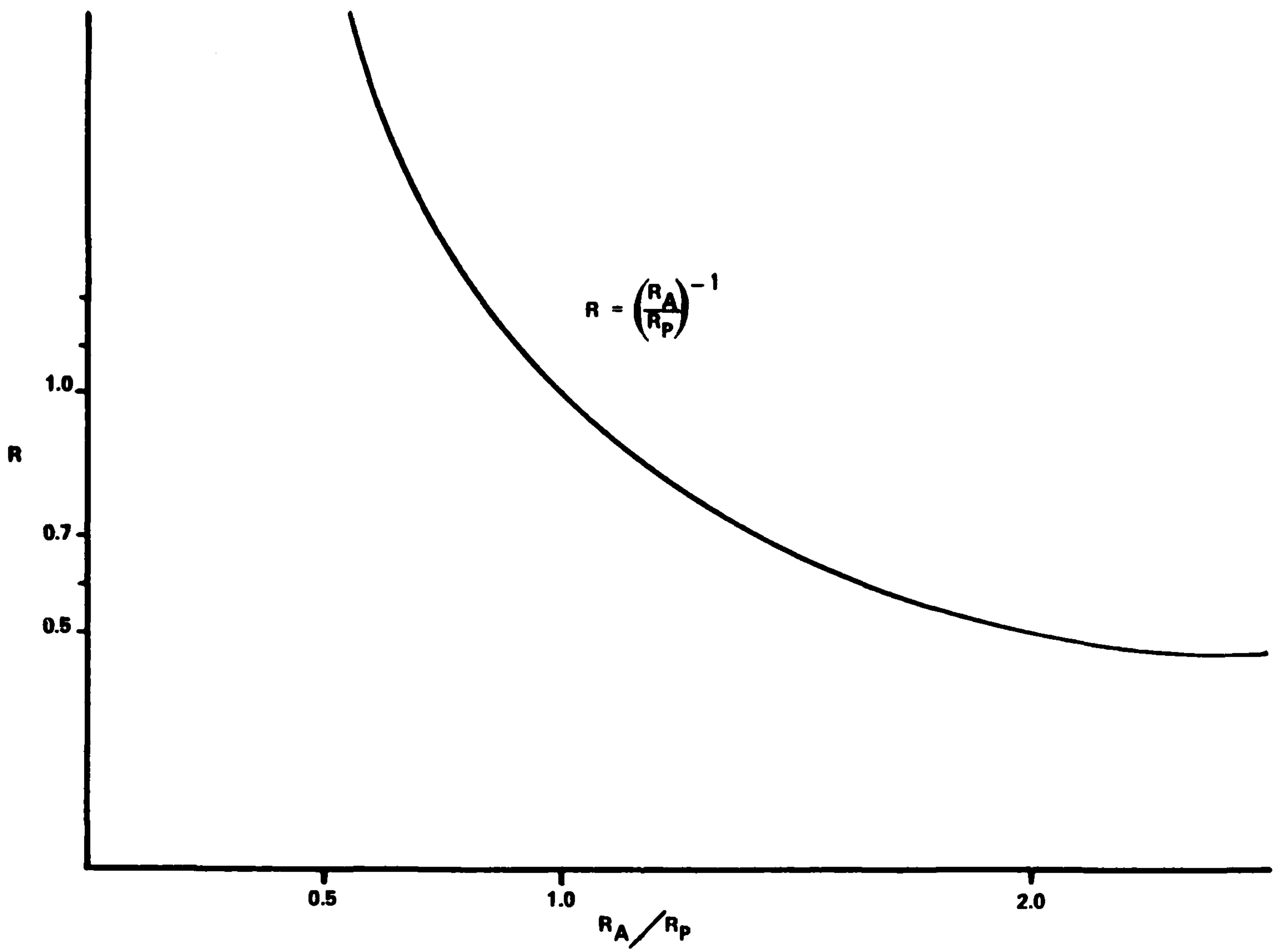


Figure F-1. Generic Efficiency Factor (R)

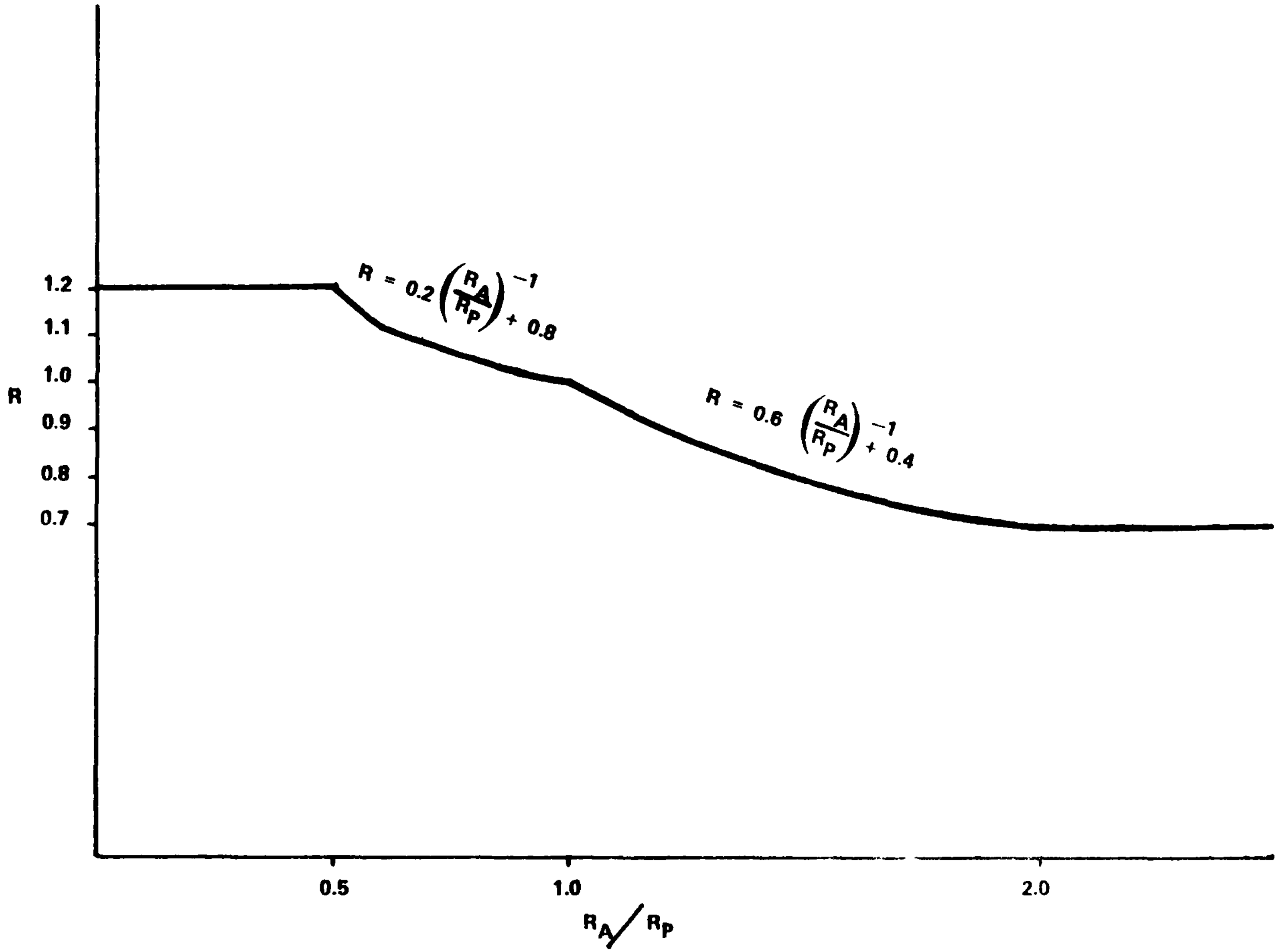


Figure F-2. TPMS Efficiency Factor (R) as a Function of R_A/R_P

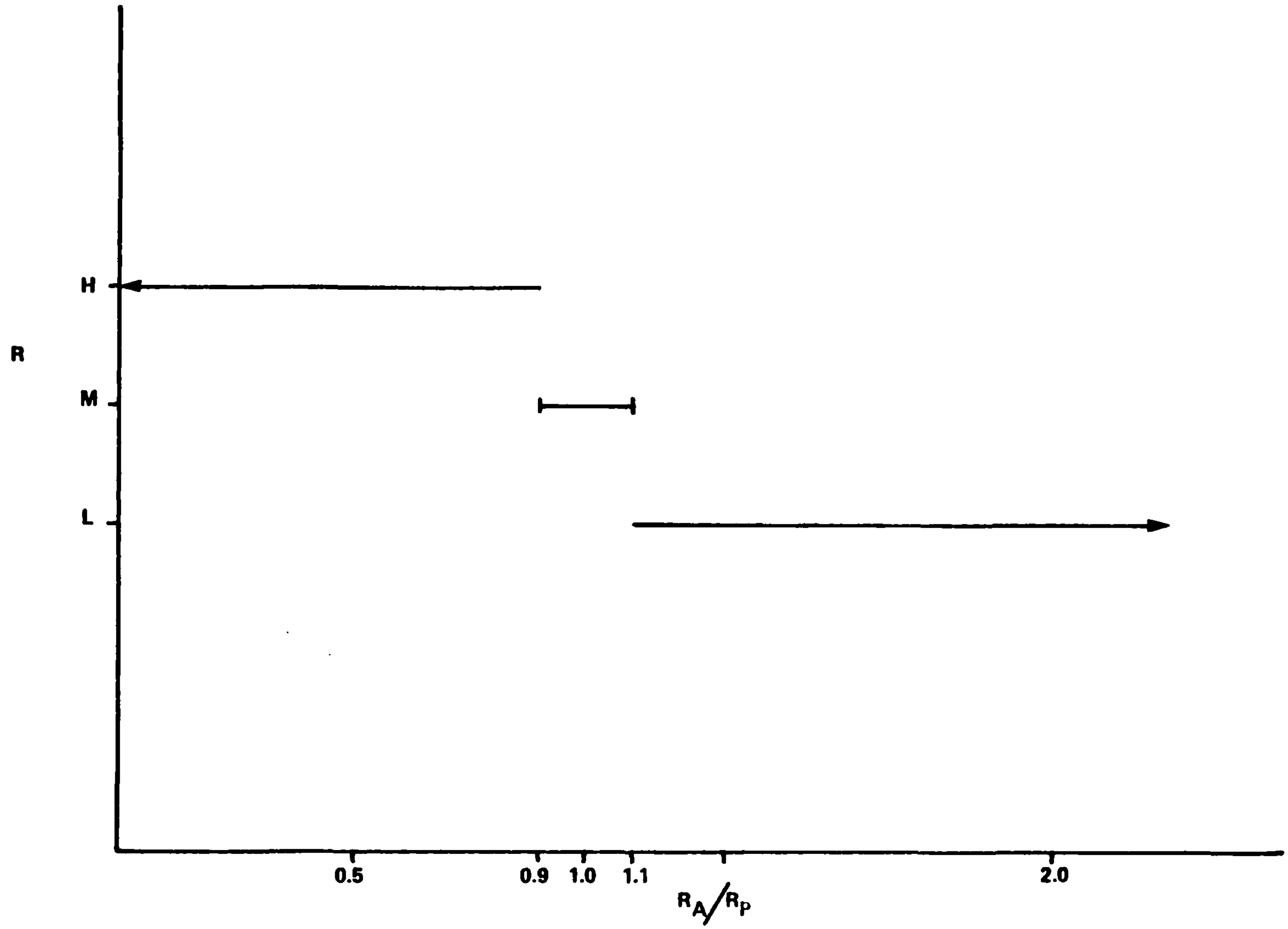


Figure F-3. Alternative Manpower Efficiency Factor (R)

b. Quality

The alternative scoring procedure requires that quality also be ordinally scaled with three categories: high (excellent), medium (okay), low (substantial revision required to release). This categorical scoring would produce the desired discrimination among projects and would be much more realistic in terms of expectations from PRB members. The potential for placing a project into the wrong category would be considerably less than the potential error in the current grading scheme.

c. Timeliness

This factor would also be scaled ordinally with the following three categories: high (early), medium (on time), low (late). Figure F-4 shows the graph for this factor.

3. COMBINING THE FACTORS

a. The three distinct levels or categories for each of the three factors results in a total of twenty-seven distinct possible results for any project. These outcomes are represented with the tree diagram in Figure F-5. Ranking of the projects on a productivity scale can now be accomplished by sorting the various outcomes according to the importance attached to the level of each factor. Numerical values could then be assigned to the sorted categories. This would allow for weighted (by manpower) averages of project scores to be computed for organizational units.

b. Table F-1 gives one possible sort of categories, assuming that each factor is of equal importance and that the levels across factors are comparable. In this scoring scheme there are seven distinct categories. If this scoring scheme were applied to the first year results, the resulting distribution is shown in figure F-6.

4. ALTERNATIVE SCORING PROCEDURE FLEXIBILITY

a. In addition to its simplicity, the alternative scoring procedure is extremely flexible. For example, if overestimation of R_p is not sufficiently discouraged, the computation of R could be altered to include a penalty window for overestimation. Figure F-7 shows two methods by which this could be done. If a project has actual manpower between 50 and 90 percent of planned manpower, an R value of "medium" could be assigned (upper dotted line), or an R value of "low" could be assigned (lower dotted line). Using the latter choice, the projects for the first scoring period were given scores by the alternative procedure. The results are shown by the histogram in figure F-8. This histogram can be compared to the one in figure F-6. Notice that discouraging or penalizing overestimating results in a more symmetrical distribution.

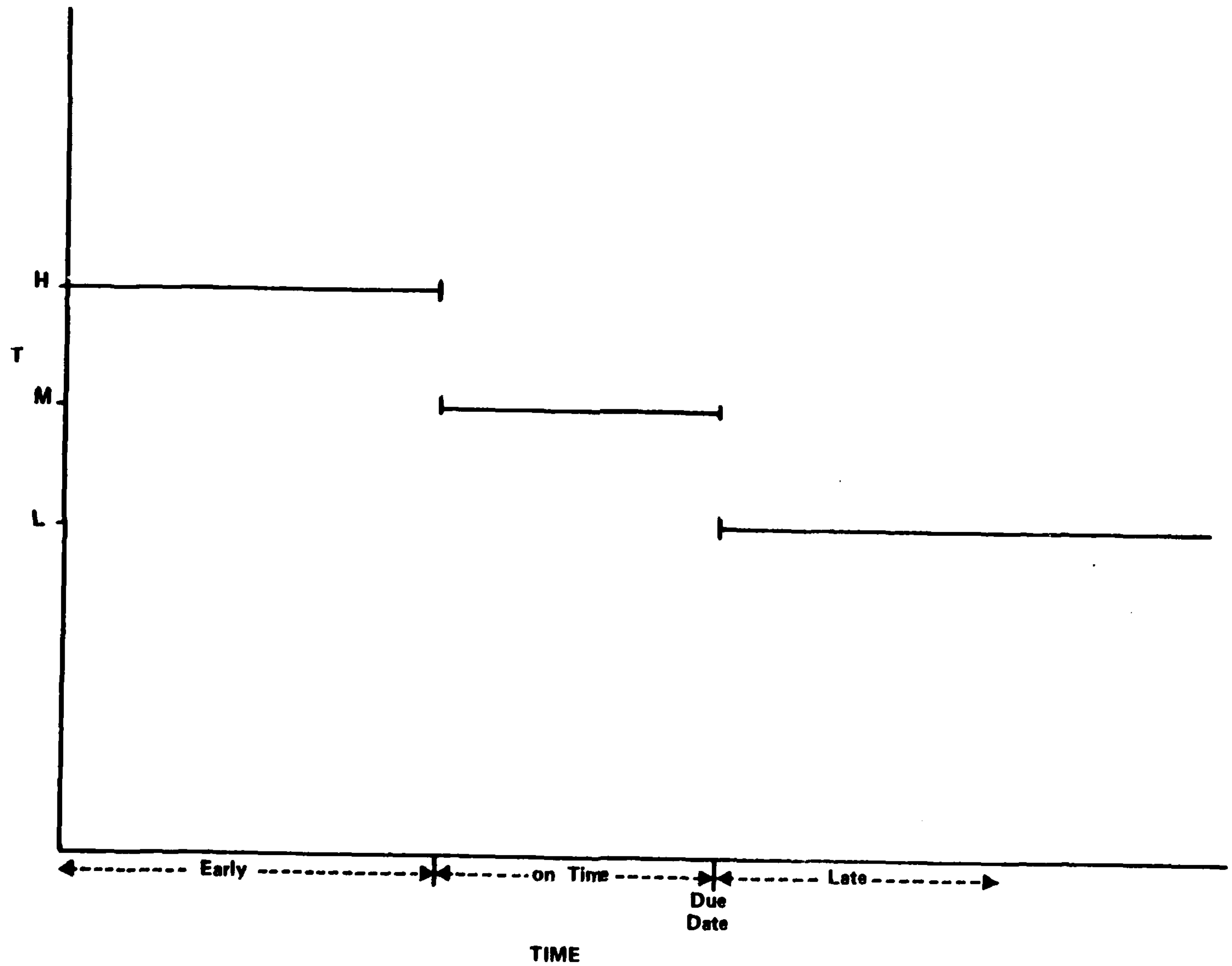


Figure F-4. Alternative Timeliness Factor (T)

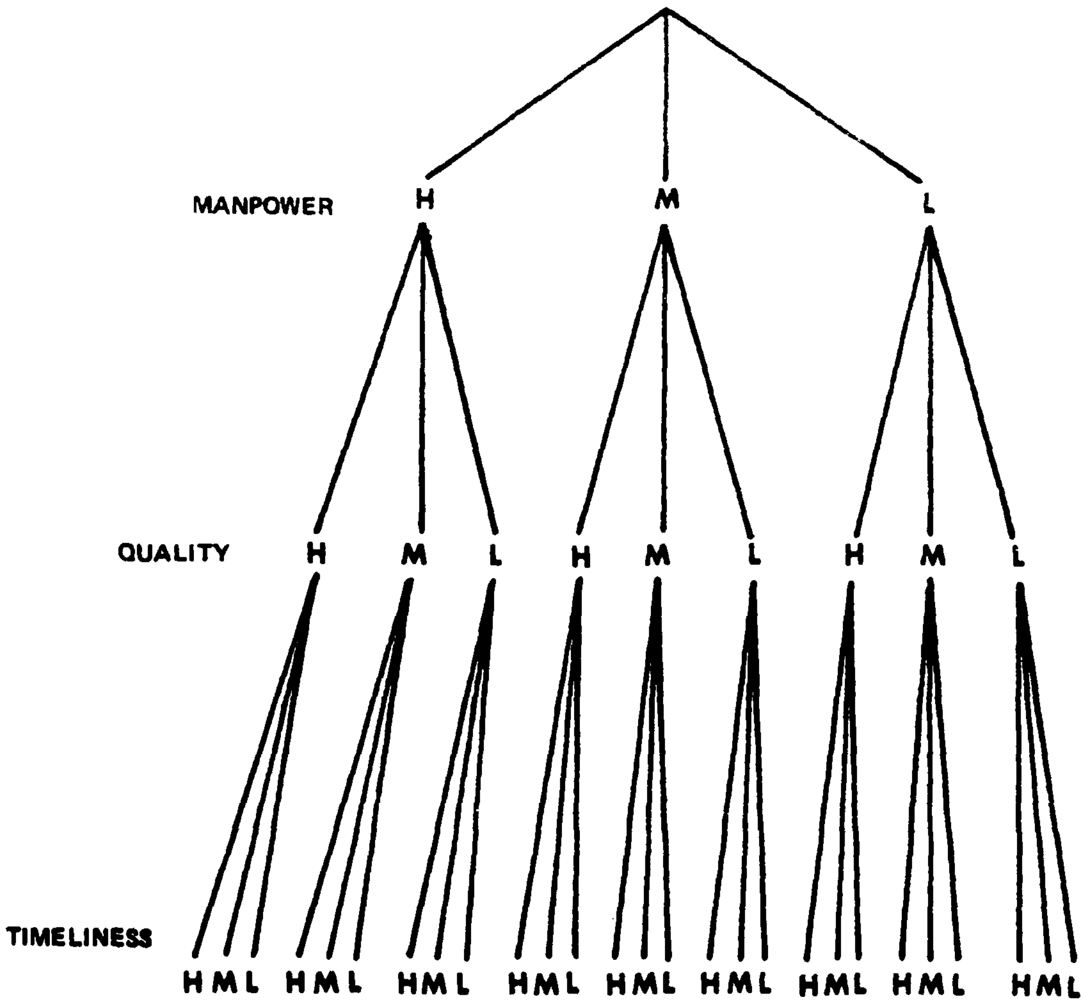


Figure 7-5. Scoring Outcomes

TABLE F-1

A RANKING OF SCORING CATEGORIES

CATEGORY RANKING	OUTCOME CATEGORY
7	(H, H, H)
6	(H, H, M), (H, M, H), (M, H, H)
5	(H, M, M), (M, H, M), (M, M, H) (H, H, L), (H, L, H), (L, H, H)
4	(M, M, M) (H, L, M) (H, M, L) (L, H, M) (M, H, L) (M, L, H) (L, M, H)
3	(L, M, M) (M, L, M) (M, M, L) (L, L, H) (L, H, L) (H, L, L)
2	(M, L, L) (L, M, L) (L, L, M)
1	(L, L, L)

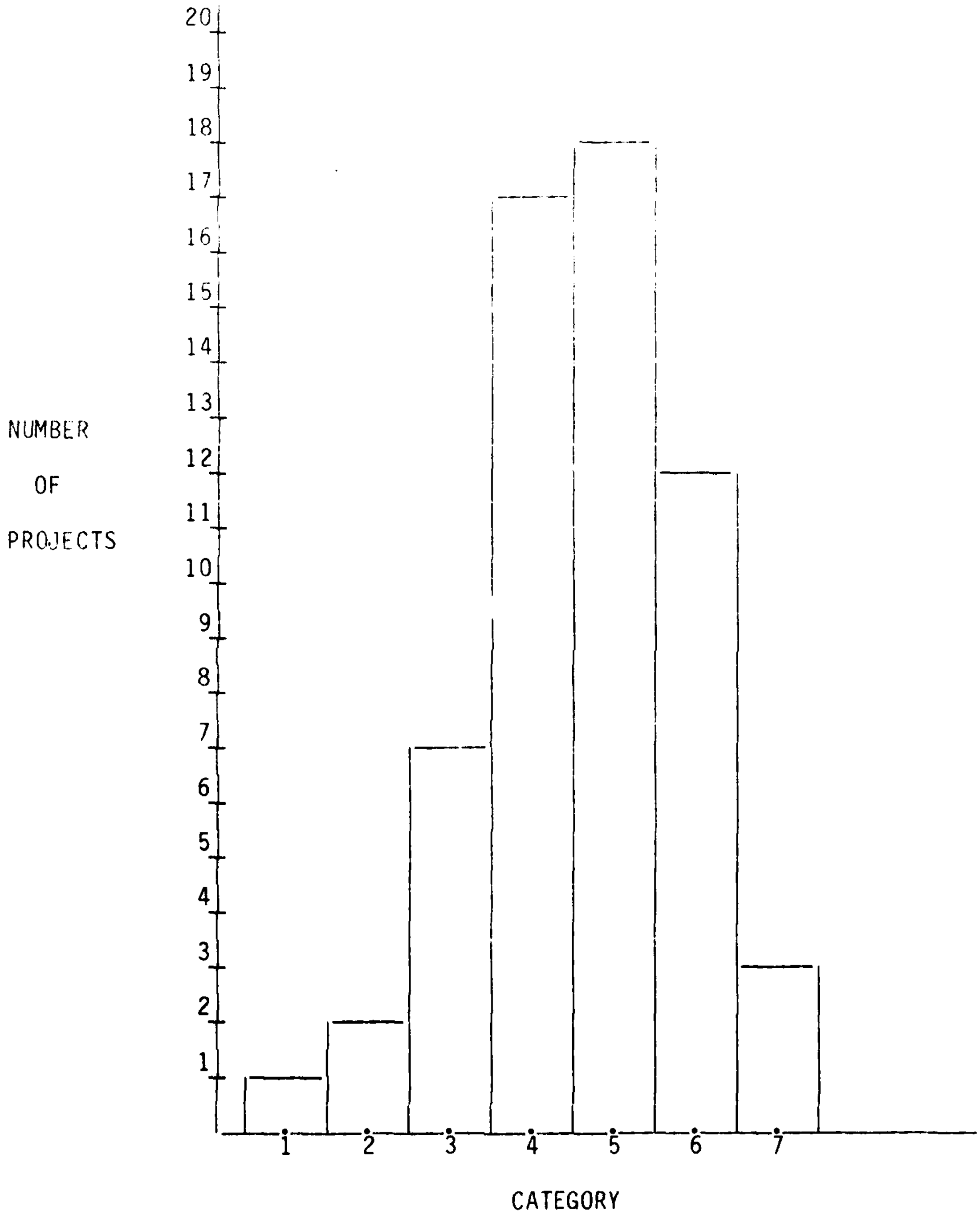


Figure F-6. First Year Project Scores Alternative Scoring Procedure

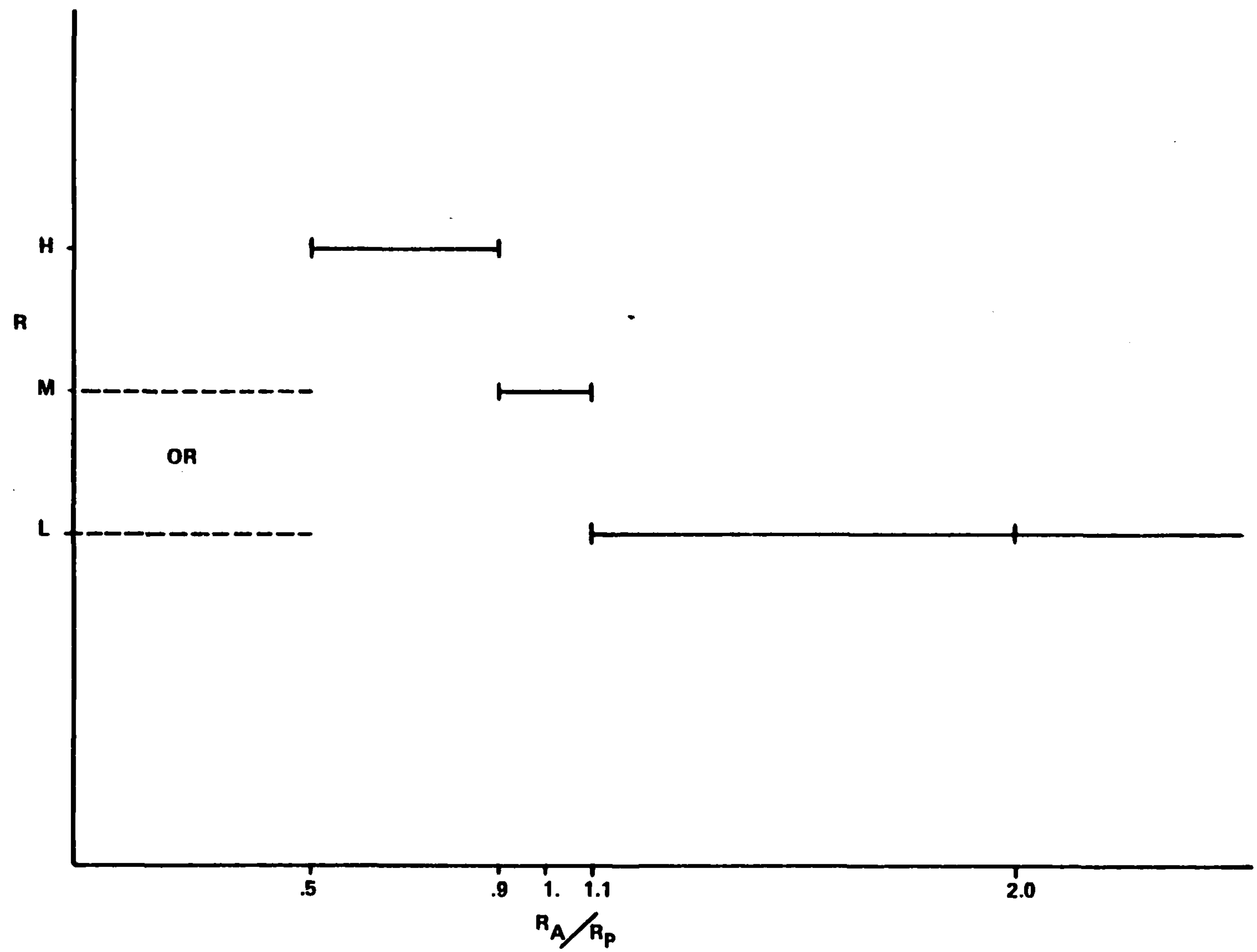


Figure F-7. Alternative Manpower Efficiency Factor with Penalty for Overestimation

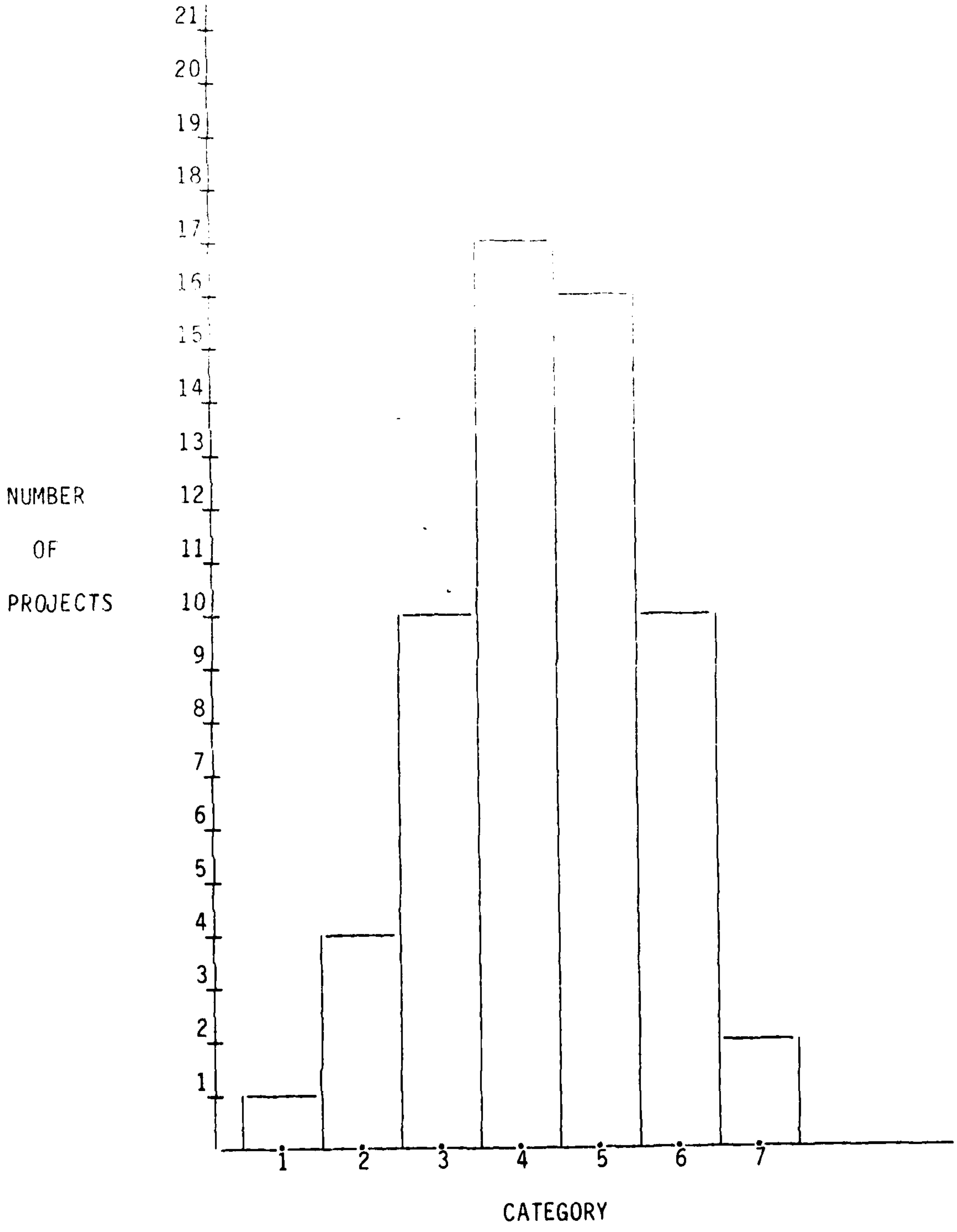


Figure F-8. Alternative Scoring Procedure with Penalty for Overestimation of Manpower Applied to First Year Project Scores

b. As another example of the flexibility of the alternative system, suppose that it is decided that the timeliness factor should be scored: high-on time, medium-late, and low-very late. The levels of this factor probably should not be treated as comparable to those of the other two, and a new ranking of the 27 scoring outcomes of figure F-5 might be required. Table F-2 gives one of many possible ranking schemes.

TABLE F-2

A RANKING OF SCORING CATEGORIES

TIMELINESS ALTERED

CATEGORY RANKING	OUTCOME CATEGORY
9	(H, H, H)
8	(H, M, H) (M, H, H)
7	(H, H, M) (L, H, H) (H, L, H) (M, M, H)
6	(H, M, M) (M, H, M) (M, L, H) (L, M, H)
5	(H, H, L) (L, L, H) (M, M, M) (H, L, M) (L, H, M)
4	(H, M, L) (M, L, M) (M, H, L) (L, M, M)
3	(H, L, L) (L, H, L) (M, M, L) (L, L, M)
2	(M, L, L) (L, M, L)
1	(L, L, L)

c. The use of three levels for each of the factors seems to be simple to use yet provides enough discrimination among the projects; however, more or fewer levels might also be considered. For example, the timeliness factor might be altered to have only two levels: medium-on time and low-late. If the other factors remain the same, the tree diagram of figure F-9 shows the 18 possible outcomes for this scheme and table F-3 gives one possible ranking of the outcomes.

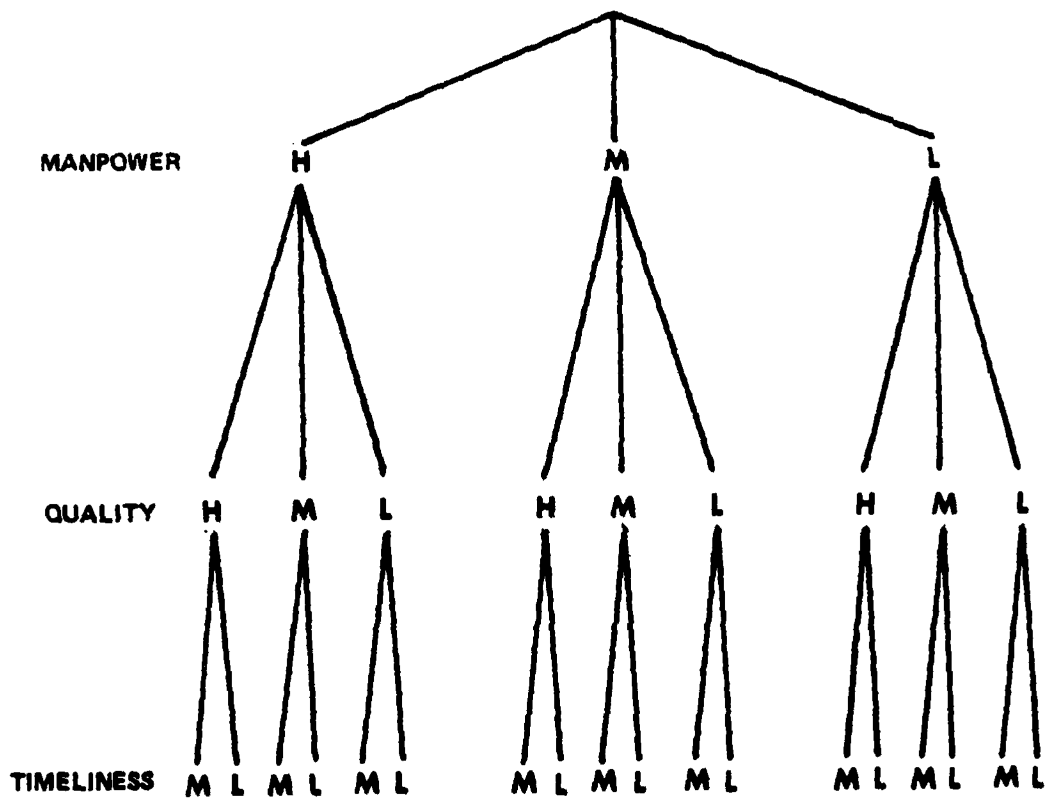


Figure F-9. Scoring Outcomes with Two Levels for Timeliness

TABLE F-3

A RANKING OF SCORING CATEGORIES

TWO LEVELS FOR TIMELINESS

CATEGORY RANKING	OUTCOME CATEGORY
6	(H, H, M)
5	(H, H, L) (H, M, M) (M, H, M)
4	(H, M, L) (H, L, M) (M, H, L) (M, M, M) (L, H, M)
3	(H, L, L) (L, H, L) (M, M, L) (L, M, M) (M, L, M)
2	(M, L, L) (L, M, L) (L, L, M)
1	(L, L, L)

APPENDIX G

BIBLIOGRAPHY

BIBLIOGRAPHY

GENERAL

Conte, S. D. and De Boor, C., Elementary Numerical Analysis: An Algorithmic Approach, 3rd Edition, McGraw-Hill, New York, 1980.

Gage, N. L. and Berliner, D. C., Educational Psychology, Rand McNally, Chicago, 1975.

Glass, G. V. and Stanley, J. L., Statistical Methods in Education and Psychology, Prentice-Hall, Englewood Cliffs, NJ, 1970.

Goodman, A. W., Analytic Geometry and the Calculus, 3rd Edition, MacMillan, New York, 1974.

Mehrens, W. A. and Lehman, I. J., Measurement and Evaluation in Education and Psychology, Holt, Rinehart & Winston, New York, 1973.

Siegel, S., Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill, New York, 1956.

TPMS

Christ, R. E., The Initial Development and Evaluation of An Algorithm for Monitoring Manpower Requirements and the Relative Difficulty of Projects Within TRASANA, US Army TRADOC Systems Analysis Activity (TRASANA), October 1981.

Crum, J., Report on the Evaluation of the Productivity Measurement System Developed by TRADOC Systems Analysis Activity (TRASANA). Personal communication to Leon F. Goode, Jr., Deputy Director for Technical Operations, TRASANA, 25 July 1980.

Dooley, D., TPMS Evaluation Summary, Memorandum for Record to Leon F. Goode, Jr., Deputy Director for Technical Operations, TRASANA, 21 July 1980.

Goode, L. F., Jr. and Meier, D. H., Productivity Measurement for Thinkers -An Overview of the TRASANA technical staff Productivity Measurement System, Bureaucrat, April 1981.

TRADOC OE Team Memorandum for Director, Organizational Effectiveness Visit, 9-12 February 1981, TRASANA, 24 February 1981.

TRASANA Memorandum 5-4, TRASANA Technical Staff Productivity Measurement System (TPMS)-Test, Department of the Army, TRADOC Systems Analysis Activity (TRASANA), August 1980.

TRASANA Memorandum 5-5, Quality Assurance of TRASANA Technical Products, Department of the Army, TRADOC Systems Analysis Activity (TRASANA), May 1981.

TRASANA Technical Documentation 37-80, Productivity Measurement System Design and Development, Department of the Army, TRADOC System Analysis Activity (TRASANA), December 1980.

Schainblatt, A. H., Measuring the Productivity of Scientists and Engineers in R&D: A State of the Practice Review, The Urban Institute, May 1981.

R&D PRODUCTIVITY MEASUREMENT

Clark, L. C. and Guba, E. G., A Study of Teacher Education Institutions as Innovators, Knowledge Producers, and Change Agencies, Indiana University, April 1977.

Corbin, L. J., Productivity and Job Satisfaction in Research and Development: Associated Individual and Supervisory Variables, Thesis, Air Force Institute of Technology, September 1977, AD A045981.

Edwards, S. A. and McCarrey, M. W., Measuring the performance of researchers, Research Management, January 1973, pp. 34-41.

Katzen, R., Measuring the productivity of engineers, Chemical Engineering Progress, April 1975, pp. 24-28.

Koser, M. C., Quantitative Scientists/Engineers Productivity and Some Associated Individual and Organizational Variables, Thesis, Air Force Institute of Technology, December 1976, AD A035752.

Measuring engineering efficiency, Chemical Engineering, 24 December 1962.

Newburn, R. M., Measuring productivity in organizations with unquantifiable end-products, Personnel Journal, September 1972, pp. 655-657.

R&D Productivity, Hughes Aircraft Co., 2nd edition, 1978.

Stevens, A. E., Rewards in Air Force R&D: An Analysis of Desirability, Perception and Association with the Productivity of Scientists/Engineers, Thesis, Air Force Institute of Technology, December 1976, AD AO 36462.

Technical Report 74-3, Research and Development Professionals: An Examination of the Personal Value System and Operative Goals of the Scientists, Engineers and Managers in One Government R&D Organization, Department of the Air Force, Air Force Institute of Technology, September 1974, AD 786877.

Technical Report 78-1, Productivity in the Air Force Weapons Laboratory: Measurement and Prediction, Department of the Air Force, Air Force Institute of Technology, April 1978, AD A052897.

Technical Research Report 1149, Motivational Factors Affecting Army Research and Development Personnel, Department of the Army, Personnel Research Office, May 1966, AD 640390.

PERFORMANCE APPRAISAL

Blose, L. E. and Stansel, M. J., Management by Objectives At A Research, Development, Test, and Evaluation Activity, Thesis, Naval Postgraduate School, June 1975, AD A013627.

Decotiis, T. and Petit, A., The performance appraisal process: a model and some testable propositions, Academy of Management Review, July 1978, pp. 635-645.

Gallagher, M. C., More bias in performance evaluation, Personnel, July-August 1978, pp. 35-40.

Lodahl, T. M. and Kejner, M., The definition and measurement of job involvement, Journal of Applied Psychology, 49(1), 1965, pp. 24-33.

McGregory, D., An uneasy look at performance appraisals, Harvard Business Review, September-October 1972, pp. 133-138.

Wallace, L., Nonevaluative approaches to performance appraisals, Supervisory Management, March 1978, pp. 2-9.

PEER RATINGS

Booker, G. S. and Miller, R. W., A closer look at peer ratings, Personnel, January-February 1966, pp. 42-47.

Borman, W. C., The rating of individuals in organizations: An alternate approach, Organization Behavior and Human Development, 12, 1974, pp. 105-124.

Fisk, D. W. and Cox, J. A., Jr., The Consistency of ratings by peers, Journal of Applied Psychology, 44(1), 1960, pp. 11-17.

Lewin, A. Y., Dubno, P. and Akula, W. G., Face to Face interaction in the peer nomination process, Journal of Applied Psychology, 55(I), 1971, pp, 495-497.

Technical Report 79-9, Peer Rating Research: Annotated Bibliography, Department of the Air Force, Human Resources Laboratory, June 1979, AD A071409.

Technical Report 79-31, Peer and Supervisory Ratings of Research Scientists, Department of the Navy, Naval Personnel and Development Center, September 1979, AD A074842.

Waters, L. K. and Waters, C. W., Peer nominations as predictors of short-term sales performance, Journal of Applied Psychology 54(1), 1970, pp. 42-44.

END

FILMED

1-83